

Universidad Carlos III de Madrid

Escuela Politécnica Superior

Ingeniería en Informática



**INTEGRACIÓN DE UN CORRECTOR DE NOMBRES PROPIOS
A PARTIR DE LA SALIDA DE UN RECONOCEDOR DE VOZ**

Proyecto de Fin de Carrera

Autor:

Jose María Chumo Mata

Tutora:

Paloma Martínez Fernández

Septiembre 2013

A mi familia, por algunas cosas.
A mis amigos, por otras cosas.

Abstract

English As computers are becoming more pervasive, engineers and designers are pushing the boundaries in human-computer interaction. One of latest trends is the automatic speech recognition. It allows users to utter complex commands with their voices. This Bachelor's Degree Thesis explores the possibilities of fusing two existing technologies (speech recognition and text correction) to create a web service and provide an API for general usage. As providing general correction capabilities is a very hard task, this project has been focused on a specific domain. In this case, soccer players.

Español Mientras el uso de ordenadores está cada vez más generalizado, ingenieros y diseñadores están ampliando los límites en la interacción humano-ordenador. Una de las tendencias que ha surgido con fuerza en los últimos años es el reconocimiento automático de voz, que permite a los usuarios dictar órdenes complejas con la vez. Este Trabajo de Fin Carrera explora la posibilidad de fusionar dos tecnologías existentes (el reconocimiento de voz y la corrección de texto) para crear un servicio web que proporcione un API para el uso general. Dado que proporcionar una corrección general es una tarea muy complicada, este proyecto se ha centrado en un dominio específico. En este caso, han sido jugadores de fútbol.

Keywords *Automatic Speech Recognition (ASR), ASR Evaluation, Audio Transcription, Speech Correction.*

Índice general

Índice de figuras	VII
Índice de tablas	IX
1. Introducción y motivación	3
1.1. Introducción	3
1.2. Motivación	3
1.3. Estructura de la documento	4
2. Estado del arte	5
2.1. Procesamiento de voz en móviles	5
2.1.1. Siri	5
2.1.2. S Voice	7
2.1.3. Nina	8
2.1.4. Sherpa	8
2.1.5. Ask Ziggy y Translingo	9
2.2. Sistema de corrección	10
3. Diseño e implementación	13
3.1. Servicio de transcripción	13
3.1.1. Integración	14
3.2. Configuración final del sistema	16
4. Pruebas y evaluación	19
4.1. Pruebas	19
4.2. Evaluación	20
5. Conclusiones	23
Bibliografía	25

A. Requisitos	29
A.1. Requisitos de Software	29
A.2. Requisitos de Hardware	30
B. Planificación y presupuesto	33
B.1. Fases del proyecto	33
B.2. Gastos de materiales	35
B.3. Gastos de personal	35
B.4. Gastos totales	36
C. Manual	37
C.1. Servicio web	37
C.1.1. Componentes del servicio	37
C.1.2. API del servicio	40
C.1.3. Interfaz web	41
C.2. Instalación de Xen	42
C.3. Servidor web	48
D. Posibles mejoras	53
E. Salida de las pruebas	57
E.1. Resultados del servicio web	57
E.2. Resultado de la evaluación	82
Índice alfabético	90

Índice de figuras

2.1. Logo de Siri	6
2.2. Ejemplos de uso de Siri	7
2.3. Logo de Samsung S Voice	7
2.4. Logo de Nina	8
2.5. Logo de Sherpa	8
2.7. Ejemplos de uso de Ask Ziggy y Translingo	9
2.6. Logo de Ask Ziggy	9
2.8. Arquitectura del sistema de corrección	11
3.1. Flujo diagrama de flujo del servicio de transcripción	14
3.2. Diagrama de flujo de la aplicación final	16
3.3. Capas del sistema final	17
B.1. Planificación del proyecto	34
C.1. Diagrama de flujo del transcriptor	39
C.2. Interfaz web del sistema	41
C.3. Iconos de estado de la grabación	42
C.4. Fragmento de la salida de <code>phpinfo</code>	50
C.5. Contenido del directorio del servicio	51

Índice de tablas

2.1. Ejemplo de entidades mal reconocidas	10
2.2. Patrones de extracción	10
2.3. Representaciones fonéticas para el español	12
4.1. Resultados obtenidos en las pruebas	20
4.2. Resultados de la evaluación con Scite	21
A.1. Definición de la tabla de requisitos de software	29
A.2. Requisito de software 1: sistema operativo	30
A.3. Requisito de software 2: Windows Speech Recognizer	30
A.4. Requisito de software 3: Java Runtime Environment 6	31
A.5. Requisito de software 4: Servidor web	31
A.6. Requisito de hardware 1: CPU	31
A.7. Requisito de hardware 2: RAM	32
A.8. Requisito de hardware 3: espacio en disco	32
B.1. Material expenses	35
B.2. Presupuesto para personal	35
B.3. Presupuesto total	36

Listings

C.1. XSD para errores	40
C.2. XSD para resultados	40
C.3. Fichero de arranque personalizado	44
C.4. Fichero de configuración para la máquina virtual	46
D.1. Registro de comandos de transcripción	54
E.1. Salida para Andrés Iniesta (voz femenina)	57
E.2. Salida para Cristiano Ronaldo (voz femenina)	58
E.3. Salida para David Villa (voz femenina)	60
E.4. Salida para Iker Casillas (voz femenina)	61
E.5. Salida para Lass Diarra (voz femenina)	62
E.6. Salida para Müller (voz femenina)	64
E.7. Salida para Özil (voz femenina)	65
E.8. Salida para Rosenberg (voz femenina)	67
E.9. Salida para Sergio Ramos (voz femenina)	68
E.10. Salida para Andrés Iniesta (voz masculina)	69
E.11. Salida para Cristiano Ronaldo (voz masculina)	71
E.12. Salida para David Villa (voz masculina)	72
E.13. Salida para Iker Casillas (voz masculina)	74
E.14. Salida para Lass Diarra (voz masculina)	75
E.15. Salida para Müller (voz masculina)	76
E.16. Salida para Özil (voz masculina)	78
E.17. Salida para Rosenberg (voz masculina)	79
E.18. Salida para Sergio Ramos (voz masculina)	81
E.19. TRN de referencia para Scite	82
E.20. TRN de voz femenina sin corregir	82
E.21. Resultado para voz femenina sin corregir	83
E.22. TRN de voz femenina corregido	84
E.23. Resultado para voz femenina corregida	84
E.24. TRN de voz masculino sin corregir	86
E.25. Resultado para voz masculina sin corregir	86
E.26. TRN de voz masculino corregido	88

E.27.Resultado para voz masculina corregida	88
---	----

Capítulo 1

Introducción y motivación

1.1. Introducción

Los sistemas de información se han convertido en un elemento ubicuo en nuestra vida diaria. Actualmente la mayoría de los usuarios interactúan con dichos sistemas mediante periféricos como el teclado y el ratón, o por medio de pantallas táctiles. Sin embargo, en el últimos años se está potenciando el uso de comandos de voz para la interacción con dispositivos electrónicos. Para poder interpretar los comandos de voz proporcionados por los usuarios se usan transcriptores de voz, un tipo de software que permite obtener una representación textual de las locuciones de los usuarios. Como con muchos otros tipos de software, existe una gran cantidad tipos e implementaciones de transcriptores. En algunos sistemas se opta por desarrollar un sistema de transcripción propio mientras que en otros se opta por usar transcriptores existentes, ya sean libres o privativos.

1.2. Motivación

En ocasiones, las transcripciones realizadas contienen errores y no consiguen interpretar correctamente las palabras de los usuarios, ya sea por el idioma, por el timbre y el tono de la voz, por la pronunciación o por otros factores. Los errores producidos varían de un transcriptor a otro por lo que su corrección no es problema fácil de atajar. Una de las posibles soluciones pasa por usar varios sistemas de transcripción, lo cual implica, entre otras cosas, tener que adaptar el sistema cliente al Application Programming Interface (API) del transcriptor cada vez que se deba agregar uno nuevo. Si se tiene en cuenta que cada vez más proyectos hacen uso de esta tecnología, nos damos cuenta de que este proceso puede ser muy caro en términos tempora-

les y económicos si se realiza de forma independiente. El objetivo principal que se pretende alcanzar con este proyecto es conseguir un sistema capaz de realizar transcripciones a partir de entradas de audio y de corregir los posibles errores. Se pretende que el software obtenido sea capaz de integrar varios sistemas de transcripción distintos para poder usar el mejor resultado en distintas situaciones.

Dado que este es un proceso que posiblemente vayan a llevar a cabo muchos otros desarrolladores se pretende que el sistema resultante ofrezca un [API](#) que facilite su integración en nuevos programas. Uno de los aspectos a tener en cuenta es que los dispositivos que harán uso de esta tecnología pueden estar limitados en cuanto a la capacidad de almacenamiento o estar desarrollados en lenguajes distintos y difíciles de integrar con el usado para el desarrollo de este proyecto. Una posible solución es diseñar el software como una plataforma Software as a Service (SaaS).

1.3. Estructura de la documento

Este documento está dividido en cuatro partes, siendo esta la primera. En esta parte se expresan las motivaciones que han impulsado el proyecto y los objetivos que se espera conseguir.

La segunda parte describe el estado del arte, es decir, las tecnologías existentes en las que se basa el proyecto o que cumplen funcionalidades parecidas a las que se quieren conseguir. En este caso se han descrito varias aplicaciones móviles de asistentes personales y el trabajo de fin de máster usado en la implementación de este proyecto.

En la tercera parte se cubre la fase de desarrollo del proyecto. En ella se detallan los pasos llevados a cabo durante las distintas fases de desarrollo así como las herramientas empleadas. También se explican las decisiones tomadas y el razonamiento detrás de cada una de ellas.

La cuarta parte de este documento está compuesto por los anexos. Contienen los manuales para la configuración y uso de la aplicación, las propuestas de mejoras para la misma y los ficheros en formato texto generados por las aplicaciones durante la fase de pruebas.

Este documento también contiene la bibliografía consultada, un índice alfabético, una lista de acrónimos y definiciones de términos usados a lo largo del documento.

Capítulo 2

Estado del arte

En este capítulo se listarán y describirán algunas aplicaciones que hacen uso de servicios web o plataformas [SaaS](#) que proporcionan una funcionalidad similar a la que se pretende conseguir. También se describirán algunos de los componentes empleados en el desarrollo de la solución propuesta en este proyecto.

2.1. Procesamiento de voz en móviles

En los últimos años se ha producido un incremento en el uso de dispositivos móviles inteligentes, con capacidades de procesamiento cada vez mayor y con la posibilidad de conectarse a Internet desde casi cualquier parte. Eso ha propiciado la aparición de sistemas pensados especialmente para hacer uso de dicha conexión y servir como herramientas de ayuda en una gran variedad de situaciones, como pueden ser los asistentes personales inteligentes.

Un asistente personal inteligente es un software que puede llevar a cabo varias tareas teniendo en cuenta órdenes indicadas por el usuario la localización e información de otras fuentes (servicio meteorológico, servicios de noticias, informes de tráfico...). Entre algunas de las funciones que desempeñan se encuentran el envío de mensajes, planificación de citas, búsquedas en internet y reservas en distintos servicios.

En esta sección se enumeran y describen algunas de las alternativas disponibles en los distintos sistemas móviles.

2.1.1. Siri

Siri, una aplicación incluida desde el año 2011 en iOS, es un asistente personal capaz de interactuar con los usuarios mediante el lenguaje natural

en lugar de mediante comandos de voz preestablecidos. El sistema es capaz de tener en cuenta las interacciones inmediatamente anteriores para proporcionar las respuestas a las consultas realizadas pedidas, de forma similar a la que la haría un humano, proporcionando así una sensación de conservación. Su interfaz consiste en una vista similar a un chat, donde se muestra la transcripción de la consultas formuladas y las respuestas proporcionadas por el sistema, que también son locutadas.



Figura 2.1: Logo de Siri

Aparte de desempeñar funciones de asistente personal, también proporciona un sistema de transcripción de voz para su uso tanto en aplicaciones de Apple como en las desarrolladas por terceros. Por ejemplo, permite dictar los correos en la aplicación para el correo electrónico en lugar de usar el teclado.

El funcionamiento del asistente depende de la localización desde la que se use. La lista de países e idiomas soportados es la siguiente:

- Estados Unidos (en inglés y español)
- Reino Unido (en inglés)
- Australia (en inglés)
- Francia (en francés)
- Alemania (en alemán)
- Japón (en japonés)
- Canadá (en inglés y francés de Canadá)
- China (en mandarín)
- Hong Kong (en cantonés)
- Italia (en italiano)
- Corea (en coreano)
- México (en español)
- España (en español)
- Suiza (en francés, alemán e italiano)
- Taiwán (en mandarín)



Figura 2.2: Ejemplos de uso de Siri

Aunque el sistema se puede usar desde cualquier país y permite elegir el idioma deseado, está diseñado para reconocer los acentos y dialectos específicos de los países indicados en la anterior lista, siendo la precisión mayor cuando los usuarios son nativos. Según Apple, Siri está pensado para adaptarse al usuario por lo que el reconocimiento de acentos y dialectos debe mejorar con el uso.

2.1.2. S Voice



Figura 2.3: Logo de Samsung S Voice

Es un asistente personal desarrollado por Samsung para sistemas Android y fue presentado en mayo del 2012. El software no está disponible en el mercado de aplicaciones de Android y viene preinstalado en dispositivos Samsung. A diferencia de Siri, S Voice no realiza una locución de todos los resultados, por lo que obliga al usuario a leer de la pantalla, y es menos eficaz en la interpretación del lenguaje natural. Es capaz de realizar varias tareas, entre ellas están realizar llamadas, enviar mensajes de texto, programar alarmas y eventos en calenda-

rios, realizar búsquedas y controlar algunas funciones del dispositivo como la WiFi y el Bluetooth.

2.1.3. Nina



Figura 2.4: Logo de Nina

Nina es un producto desarrollado por Nuance. Esta aplicación no es un asistente personal que permita manejar varias funcionalidades del dispositivo mediante interacción natural, sino que se trata de un SDK para que para proporcionar algunas funcionalidades a las aplicaciones de servicios al consumidor en dispositivos móviles (bancos, aerolíneas, hoteles...). Está disponible para Android y iOS. El sistema se apoya en servicios web para realizar las transcripciones. También proporciona servicios de identi-

cación biométrica, que es uno de los campos menos explotados en el resto de sistemas. Nina está representada como una avatar que con varios estados (escuchando, despierta...) que puede ser adaptado según las necesidades de los clientes (es posible cambiar el aspecto del avatar, agregar estados...). Nuance proporciona varias voces (femeninas y masculinas) para realizar las locuciones y ofrece la posibilidad de agregar otras.

Aparte del SDK para móviles, proporciona un SDK para la integración con la lógica de negocio alojada en servidores, para para integrar las aplicaciones móviles con los servicios web. De momento el SDK sólo está disponible para empresas y no para desarrolladores particulares.

2.1.4. Sherpa



Figura 2.5: Logo de Sherpa

Sherpa es un asistente personal desarrollado por la empresa Anboto. La aplicación ha sido desarrollada para Android y realiza la mediante el uso del lenguaje natural, aunque hace uso de un lenguaje intermedio (llamado interlingua) para realizar el análisis. Entre sus principales diferencias está que permite realizar transacciones (comprar billetes, realizar transferencias bancarias y pagos mediante PayPal...) y tiene la capacidad de reproducir música en streaming. Tiene capacidad recordar las conversaciones e interacciones con los usuarios y aprender. Las respuestas de la aplicación a las consultas se realizan simultáneamente mediante locuciones de voz y texto. A fecha de hoy (febrero del 2013) la aplicación sigue en fase beta.



Figura 2.7: Ejemplos de uso de Ask Ziggy y Translingo

2.1.5. Ask Ziggy y Translingo



Figura 2.6: Logo de Ask Ziggy

Ask Ziggy y Translingo son dos aplicaciones para móviles desarrolladas por la empresa Ask Ziggy. Translingo es un sistema de traducción capaz de traducir a más de 50 idiomas. La aplicación ha sido desarrollada para iOS y es capaz de traducir lenguaje hablado, texto escrito y texto procedente de imágenes (permite seleccionar una porción de las imágenes para ser traducidas). El texto traducido se muestra por pantalla y es locutado en el idioma elegido.

La aplicación Ask Ziggy es un asistente personal desarrollado para Windows Phone. Es capaz de realizar traducciones haciendo uso de Translingo. Al igual que otros asistentes personales permite la interacción entre el usuario y el resto de aplicaciones usando el lenguaje natural.

2.2. Sistema de corrección

El sistema de corrección empleado es el implementado en el trabajo de fin de máster *Propuesta de corrección ortográfica de entidades en preguntas por voz en un sistema de QA* [19]. El objetivo del trabajo era desarrollar un sistema de corrección de entidades nombradas para un dominio concreto (en este caso noticias deportivas de fútbol) y con independencia del reconocedor utilizado.

La motivación de este trabajo partía de que los ASR producen salidas incorrectas cuando la entrada no se encuentra en su base de datos o cuando la reconoce mal. Eso, unido a que es muy costoso (y a veces imposible) para los usuarios modificar los ASR propició que se buscara una forma alternativa de mejorar la salida recibida por los usuarios de los ASR.

Consulta original	Consulta transcrita
¿Cuál fue la última película dirigida por Woody Allen ?	¿Cuál fue la última película dirigida por Raúl González ?
¿En qué equipo juega Kun Agüero ?	¿En qué equipo juega una huelga ?

Cuadro 2.1: Ejemplo de entidades mal reconocidas

El sistema propuesto usa entidades dependientes del dominio (en este caso noticias deportivas) para realizar las correcciones. Para ellos se basa en las similitudes fonéticas entre los textos reconocidos y las entidades nombradas almacenadas en la base de datos.

Como se puede ver en la figura 2.8, el sistema recibe la entrada del ASR después de que este haya transcrito la señal de voz. Después de esta fase el módulo de extracción se encarga de analizar la consulta y buscar entidades en ella. Para ello usa los patrones descritos en la tabla 2.2. Aunque esta aplicación soporta el tratamiento de consulta generales, este trabajo se ha restringido a consultas en las que sólo se utiliza la entidad a corregir.

Consulta original
¿En qué equipo juega ##JUGADOR ?
¿Quién es el máximo goleador del ##EQUIPO ?
¿Cuántos goles ha marcado ##JUGADOR en este año?
¿Cuántos penalties se pitaron en el último partido que se jugó en ##ESTADIO ?

Cuadro 2.2: Patrones de extracción

Dado que las transcripciones pueden contener errores, las comparaciones directas no son adecuadas para determinar el patrón adecuado. En su lugar,

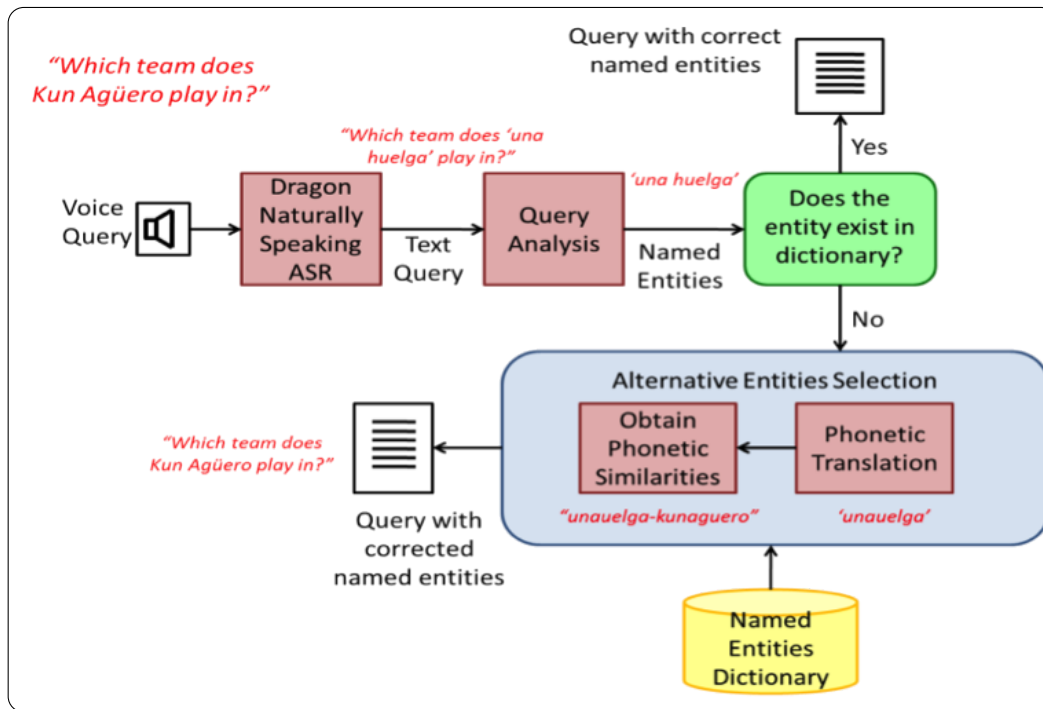


Figura 2.8: Arquitectura del sistema de corrección

se usa una bolsa de palabras¹. Para encontrar el patrón correspondiente con la consulta se cuenta el número de palabras de la consulta. Una vez determinada la regla correspondiente, se extrae la entidad por su posición en la consulta.

El siguiente paso consiste en comprobar si la entidad ha sido reconocida correctamente. Para ello se comprueba si la entidad reconocida se encuentra almacenada en el diccionario. Si se encuentra la entidad, se considera que ha sido correctamente reconocida y no se realiza la corrección. En caso de que no se encuentre en el diccionario se proporcionan alternativas a la transcripción realizada. Las reglas de la representación fonética usada para el español se pueden ver en la tabla 2.3.

Partiendo de este proyecto, se va a desarrollar un servicio web que permita realizar consultas y devuelva el resultado de las mismas con las correspondientes correcciones en caso de que sea necesario.

¹Técnica usada en procesamiento de lenguaje natural y recuperación de información. Consiste en representar un texto como una colección desordenada de palabras.

Carácter	Fonema	Carácter	Fonema	Carácter	Fonema
a	a	k	k	t	t
b	b	l	l	u	u
c	c	m	m	v	b
d	d	n	n	w	ui
e	e	ñ	N	x	ks
f	f	o	o	y	i
g	g/j	p	p	z	z
h	-	q	k	espacio	-
i	i	r	r/R		
j	j	s	s		

Cuadro 2.3: Representaciones fonéticas para el español

Capítulo 3

Diseño e implementación

La aplicación desarrollada en este proyecto consta de dos partes claramente diferenciadas: un sistema de transcripción y un sistema de corrección. En este apartado se explican las decisiones tomadas durante la fase de diseño e implementación del sistema de transcripción y su integración con el sistema de corrección descrito en el apartado [2.2](#).

3.1. Servicio de transcripción

Como se indicó en el apartado [1.2](#), uno de los objetivos es que el sistema sea fácilmente integrable en otros sistemas y aplicaciones. Ello representa un reto dada la variedad de lenguajes y sistemas operativos existentes, y plantea también la cuestión de cómo distribuir el servicio (al que posiblemente se pueden acoplar varios motores de transcripción) sin que repercuta en el tamaño de los binarios generados. Teniendo en cuenta estas cuestiones, y dada la disponibilidad de casi permanente de conexión a internet, se ha optado por ofrecer un [API](#) en forma de servicio web .

El servicio de transcripción desarrollado está pensado para realizar dos tipos de servicio: uno diseñado para transcribir clips de audio de corta duración y de no más de 30MB (enfocado a sistemas de Question Answering (QA)) y otro diseñado para clips de audio de larga duración (enfocado a la transcripción de clips de audio extraídos de videos). Ambos servicios proporcionan por defecto 2 motores de transcripción: Dragon NaturallySpeaking (DNS) y Windows Speech API (SAPI).

El servicio de transcripciones para [QA](#) es indicado para situaciones que requieran inmediatez (como en el caso de los asistentes personales inteligentes) ya que devuelve el resultado en el acto, mientras que el servicio para ficheros de audio de larga duración no proporciona respuesta inmediata. La

necesidad de inmediatez en el sistema de QA plantea dos problemas a la hora de usar DNS:

- Es necesario que la instancia esté creada antes de realizar las peticiones ya que tarda un tiempo en ser lanzada.
- El uso de hilos distintos para procesar varios ficheros simultáneamente no es posible ya que DNS crea una sola instancia del motor y se mezclan los resultados.

Para afrontar estos problemas se ha optado por separar ambos subsistemas en dos máquinas. Aunque el sistema es capaz de trabajar con varios reconocedores y se ha probado satisfactoriamente para DNS y para SAPI, éste trabajo se ha centrado en el uso de SAPI como motor de reconocimiento y en las transcripciones basadas en QA.

El diagrama de flujo mostrado en la figura 3.1 describe el funcionamiento del servicio web:

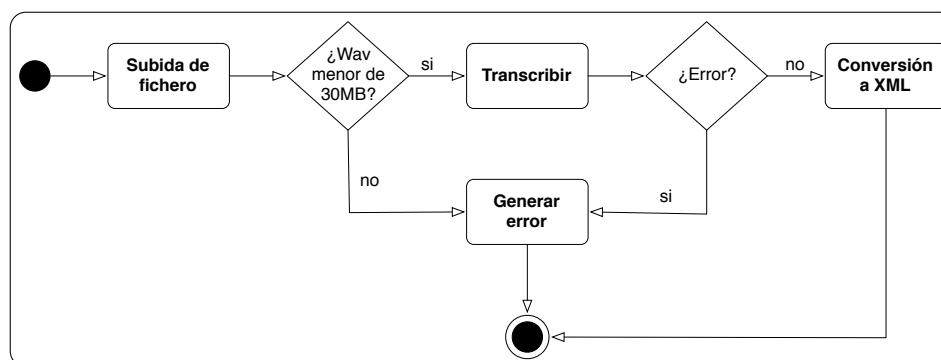


Figura 3.1: Flujo diagrama de flujo del servicio de transcripción

En la sección C.1 se describe en más profundidad la tecnología usada para implementar este módulo y las partes que lo componen.

3.1.1. Integración

Una vez diseñado el funcionamiento del sistema de transcripción se da paso a la fase de integración. El objetivo de esta fase es conectar el servicio web de transcripción con el sistema de corrección para que el resultado devuelto al cliente incluya opciones de corrección en caso de que sea necesario. Para llevar a cabo la integración se han barajado las siguientes opciones:

Monitorización del sistema de fichero Esta opción consiste en que tanto el servicio de transcripción como el de corrección monitoricen directorios predeterminados en el servidor. Las transcripciones realizadas se escribirían en el directorio monitorizado por el corrector, que al detectar el nuevo fichero lo procesaría automáticamente, escribiendo el resultado en el directorio monitorizado por el servicio de transcripción. Esta opción presenta problemas de concurrencia e implicaría muchos cambios en la implementación del corrector.

Comunicación por socket Esta aproximación implica una comunicación activa entre ambos servicios mediante un socket. Cada vez que se transcriba un fichero de audio, el servicio web le pasará al corrector la ruta en la que se haya escrito el resultado, y una vez corregida la transcripción el sistema de corrección usará el socket para comunicar localización del fichero que la contiene.

Esta opción presenta dos problemas. Por una lado se debe reservar un puerto para el socket y por otro agrega problemas de concurrencia ya que cada uno de los servicios se debe encargar de crear los hilos necesarios para la transcripción y la corrección.

Invocación como un comando Esta aproximación implica mantener el funcionamiento actual del corrector descrito en 2.2 e invocarlo desde el servicio web como un comando. De esta forma el servidor web se encarga de crear los hilos o procesos necesarios para atender las peticiones, y cada uno de esos procesos crea una máquina virtual de Java para realizar las correcciones. De esta forma se eliminan los problemas de sincronización y concurrencia a la vez que se reduce la cantidad de cambios a realizar en el corrector. La única modificación necesaria es la implementación de un parser para el formato usado en el fichero de transcripción y realizar la escritura de la corrección en XML. En este caso, el punto negativo se encuentra en el consumo de memoria, ya que se usaría una instancia de la máquina virtual de Java por cada transcripción en lugar de usar hilos.

Teniendo en cuenta los planteado anteriormente se ha optado por la tercera aproximación, ya es la que requiere menos cambios en las aplicaciones ya existentes. La figura 3.2 muestra el diagrama de flujo de la aplicación completa después de haber incluido el sistema de corrección.

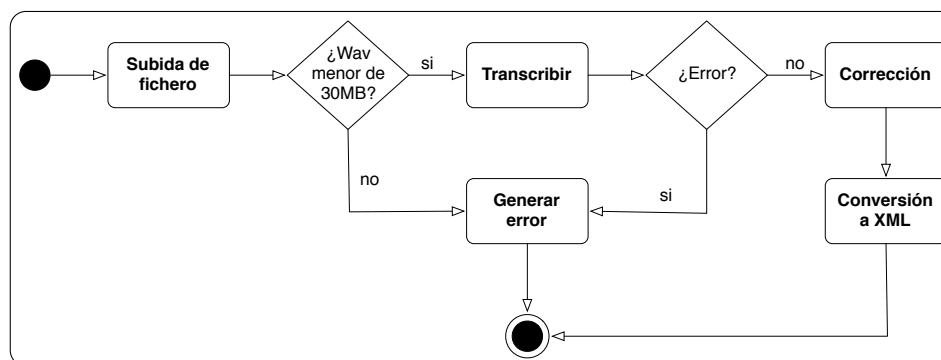


Figura 3.2: Diagrama de flujo de la aplicación final

3.2. Configuración final del sistema

Teniendo en cuenta los requisitos expuestos en [A](#), se ha elegido la siguiente configuración para implantar y probar el proyecto.

Servidor hardware: un requisito indispensable del servidor elegido es que sea capaz de soportar hipervisores por lo que se ha optado por un servidor Sun Fire X4150. En este caso está equipado con 2 procesadores Intel® Xeon® E5345 a 2.33GHz. Estos procesadores tienen 2 caches L2 de 4MB, soporte para SSE2 y permite paravirtualización. También está equipado con 8GB de RAM, que estimamos serán suficientes para alojar dos máquinas virtuales. Usar paravirtualización proporcionará flexibilidad a la hora de gestionar la instalación de las máquinas en la que se despliegue el servicio (configuración, backup) y abarata costes ya que no es necesario disponer de varias máquinas físicas. Por otro lado, la paravirtualización permite a los sistemas operativos huésped usar el hardware casi sin restricciones y sin mermar el rendimiento.

Sistema operativo base: como sistema operativo en el que instalar el hipervisor se ha escogido un Ubuntu 10.04.3 LTS de 64 bits, con la versión del kernel 2.6.32.56.

Hipervisor: para ejecutar las máquinas virtuales se hace uso de Xen 4.0.3. . El hipervisor se describe en más profundidad en la sección [C.2](#).

Sistema operativo: en las máquinas virtuales se ha instalado Windows 7. Se ha dotado a cada máquina de un disco duro de 30GB y 1GB de RAM.

Servidor web: dada la facilidad de instalación y administración se ha escogido WAMP Server 2.2 como servidor web. Esta versión incluye el servidor HTTP Apache 2.2.21 y PHP 5.3.10.

La configuración del sistema se explica en detalle en el apartado [C.3](#).

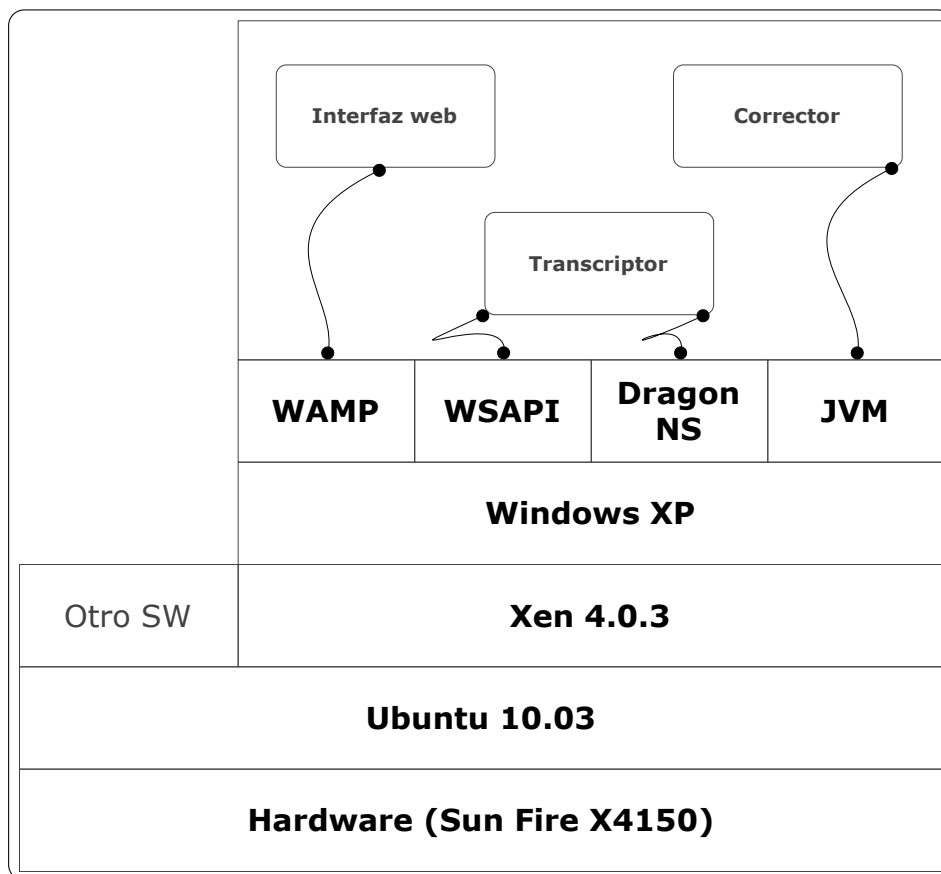


Figura 3.3: Capas del sistema final

Capítulo 4

Pruebas y evaluación

Una vez diseñado e implementado el sistema es necesario comprobar su funcionamiento mediante una serie de pruebas. En este apartado se describen las pruebas realizadas para verificar el comportamiento de la aplicación y así como el método usado para evaluar los resultados obtenidos.

4.1. Pruebas

Como ya se ha indicado anteriormente, el software desarrollado está acotado para realizar correcciones sobre un conjunto pequeño y bien definido de palabras ya que de otro modo el proyecto sería difícilmente abordable. En este caso se ha escogido como dominio los nombres de futbolistas. La base de datos de pruebas contiene las siguientes variaciones de nombres para realizar las correcciones: *Cristiano Ronaldo, Cristiano, Miguel García, Iker Casillas, Casillas, Nolito, Tiago, Jeffren, Etoó, Ronaldinho, Deco, Gerard Piqué, Forlán, Xavi, Villa, Andrés Iniesta, Iniesta, Puyol, David Villa, Edú, Lionel Messi, Leo Messi, Messi, Raúl, Müller, Andrés Palop, Palop, Kun Agüero, Sergio Ramos, Benzema, Marcelo, Carbalho, Xabi Alonso, Özil, Maxwell, Alves, Rivaldo, Alcántara, Santín, César, Soldado, Aduriz, Valdés, Di María, Canales, Dani Alves, Pepe, Rosenberg, Miguel Albiol, Xavi Hernández, Guardiola, Keita, Pedro León, Sneijder, Cesc, Aguirreche, Diego Forlán, Riki, Llorente, Aranzubía, Griezmann, Pinto, Dani Güiza, Güiza, David De Gea, De Gea, Xavi Prieto, Rafael Van der Vaart, Fernando Torres, Carles Puyol, Didier Drogba, Falcao, Van Nistelrooy, Lass Diarra.*

De los nombres presentes en la base de datos se han usado los siguientes en las pruebas: *Andrés Iniesta, David Villa, Lass Diarra, Özil, Sergio Ramos, Cristiano Ronaldo, Iker Casillas, Müller, Rosenberg.* Estos nombres han sido grabados en las mismas condiciones (entorno y hardware) tanto para voces

masculinas como para femeninas. Para la elección de los nombre se ha intentado obtener una mezcla de nombres españoles y no españoles. Los resultados obtenidos de la aplicación se pueden ver de forma resumida en la tabla 4.1, estando los ficheros completos disponibles en la sección E

nombre	reconocido	corregido	valor de mejor resultado	valor del resultado deseado
Andrés Iniesta	Andrés iniesta (f)	Andrés Iniesta	0.6666666	0.6666666
	Andrés iniesta (m)	Andrés Iniesta	0.6666666	0.6666666
Cristiano Ronaldo	Cristianos Donaldo (f)	Cristiano Ronaldo	0.8235294	0.8235294
	Cristiano Ronaldo (m)	Cristiano Ronaldo	0.8235294	0.8235294
David Villa	David villa (f)	David Villa	0.6363636	0.6363636
	David villa (m)	David Villa	0.6363636	0.6363636
Iker Casillas	Iker casillas (f)	Iker Casillas	0.7692308	0.7692308
	Iker casillas (m)	Iker Casillas	0.7692308	0.7692308
Lass Diarra	Lars diana (f)	Lass Diarra	0.45454544	0.45454544
	Lanz Viagra (m)	Lass Diarra	0.45454544	0.45454544
Müller	Molano (f)	Soldado	0.57142854	0.14285713
	Ni hule (m)	Villa	0.39999998	0.28571427
Özil	O sea (f)	Rosenberg	0.3333333	0
	O civil (m)	Özil	0.3333333	0.3333333
Rosenberg	Rusia no (f)	Cristiano	0.6666666	0.111111104
	Rozan de (m)	Llorente	0.375	0.3333333
Sergio Ramos	C en Girona los (f)	Sergio Ramos	0.4166667	0.4166667
	Sergio ramos (m)	Sergio Ramos	0.6666666	0.6666666

Cuadro 4.1: Resultados obtenidos en las pruebas

4.2. Evaluación

Para la evaluación se ha seguido el método propuesto en [6] que consiste en comparar el texto obtenido por el sistema (hipótesis) con la transcripción literal de la entrada (referencia). Dicha comparación se realiza midiendo la tasa de errores por palabra (*word error rate* o *WRE*) y la exactitud de las palabras (*word accuracy* o *WAcc*):

- **WRE:** mide el porcentaje de palabras incorrectas con respecto al número total de palabras. La fórmula es:

$$\mathbf{WRE} = \frac{ne}{pt} = \frac{ps + pi + pb}{pt} \quad (4.1)$$

donde *ne* es el número total de errores en la hipótesis y *pt* es el número total de palabras del texto de referencia. Como se puede ver en la ecuación *ne* es la suma de *ps* (sustituciones), *pi* (inserciones) y *pb* (eliminaciones). En este caso se considera que el uso incorrecto de mayúsculas es una sustitución.

- **WAcc**: mide el número total de palabras correctas con respecto al número total de palabras.

$$\mathbf{WAcc} = 1 - WER = \frac{pc}{pt} \quad (4.2)$$

donde pc es el número total de palabras correctas en la hipótesis.

Para realizar dichas mediadas usaremos el software Scrite , que forma parte del Scoring Toolkit (SCTK) desarrollado por el National Institute of Standards and Technologies (NIST). Los resultados serán agrupados por el género de la voz y diferenciando entre el resultado sin y con corrección.

En este caso se ha usado el formato trn tanto para la hipótesis como para el fichero de referencia.

Los resultados obtenidos se reflejan en la tabla 4.2. A simple vista se puede comprobar que el sistema de corrección mejora notablemente la salida obtenida. Esta conclusión es desarrollada en mayor de talle en el apartado 5.

	voz femenina		voz masculina	
	sin corrección	corregida	sin corrección	corregida
% correcto	40.0	80.0	66.7	86.7
% sustituciones	60.0	20.0	33.3	13.3
% borrados	0.0	0.0	0.0	0.0
% inserciones	26.7	0.0	20.0	0.0
% WAcc	13.3	80.0	46.7	86.7

Cuadro 4.2: Resultados de la evaluación con Scrite

Capítulo 5

Conclusiones

El objetivo del proyecto que era de integración ha sido alcanzado. Se ha conseguido aunar todas las herramientas de las que se disponía para obtener un sistema con la capacidad de transcribir y sugerir alternativas a a las transcripciones. Sin embargo todavía ha lugar para realizar varias mejoras que permitirían ofrecer un mejor servicio a los usuarios.

En cuanto a la capacidad de transcripción y corrección, como se puede ver en la tabla [4.2](#) se mejoran los resultados obtenidos por las transcripciones sin corregir, independientemente de si la voz es masculina o femenina. Consideramos que el sistema propuesto puede ser útil en entornos acotados pero no para tratar problemas generales, por lo que es necesario conocer previamente el vocabulario del entorno en el que se va a utilizar.

Bibliografía

- [1] *Proof-of-Concept for Orthographic Named Entity Correction in Spanish Voice Queries*, 2012. 2.2
- [2] Ask Ziggy. <http://www.ask-ziggy.com>. Último acceso en febrero del 2013. 2.1.5
- [3] Try Translingo. <http://www.translingoapp.com>. Último acceso en febrero del 2013. 2.1.5
- [4] Juan Castromil. Sherpa, el Siri español para Android. <http://blogs.20minutos.es/clipset/sherpa-el-siri-espanol-para-android>, 2012. Último acceso en febrero del 2013. 2.1.4
- [5] Jessica Dolcourt. Apple's Siri trounces Samsung's S Voice (hands-on). http://www.cnet.com/8301-17918_1-57458122-85/apples-siri-trounces-samsungs-s-voice-hands-on. Último acceso en febrero del 2013. 2.1.2
- [6] María González, Julián Moreno, José Luis Martínez, and Paloma Martínez. An Illustrated Methodology for Evaluating ASR Systems. 4.2
- [7] Thiago Martins. Steps to try Xen 4.0.0 Release Candidate 8 on Ubuntu Lucid 10.04 64 Bits. <http://blog.xen.org/index.php/2010/03/26/steps-to-try-xen-4-0-0-release-candidate-8-on-ubuntu-lucid-10-04-64-bits>. Último acceso en octubre de 2012. C.2
- [8] Javier Martín. El asistente de voz sherpa, gratis para los android. http://tecnologia.elpais.com/tecnologia/2012/09/28/actualidad/1348852957_020054.html, 1 octubre 2012. Último acceso en febrero del 2013. 2.1.4
- [9] Miguel Michán. Siri contra Samsung S Voice y Google Search (Jelly Bean). <http://www.applesfera.com/apple/siri-contra-samsung-s-voice-y-google-search-jelly-bean>, 2012. 2.1.2

- [10] WAMP Server. <http://www.wampserver.com>. Último acceso en octubre de 2012. 3.2
- [11] Windows 7 system requirements. <http://windows.microsoft.com/is-IS/windows7/products/system-requirements>. Último acceso en octubre de 2012. ??, ??, ??
- [12] Windows Speech API 5.4. [http://msdn.microsoft.com/en-us/library/ee125663\(v=vs.85\).aspx](http://msdn.microsoft.com/en-us/library/ee125663(v=vs.85).aspx). Último acceso en octubre de 2012. C.1.1
- [13] Windows Speech Recognition. <http://www.microsoft.com/enable/products/windowsvista/speech.aspx>. Último acceso en octubre de 2012. C.1.1
- [14] Windows Vista system requirements. <http://windows.microsoft.com/en-US/windows-vista/products/system-requirements>. Último acceso en octubre de 2012. ??, ??
- [15] Get Nina™. <http://www.nuance.com/landing-pages/enterprise/meet-nina/default.asp>. 2.1.3
- [16] Interactive product brochure :: Nina™: The Virtual Assistant for Mobile Customer Service Apps. http://www.nuance.com/landing-pages/enterprise/meet-nina/pdf/Nina_Brochure.pdf. 2.1.3
- [17] Meet Nina: The Virtual Assistant for Mobile Customer Service. <http://www.nuance.com/landing-pages/enterprise/meet-nina/pdf/NUCC1679.pdf>. 2.1.3
- [18] Julián Moreno Schneider, José Luis Martínez Fernández, and Paloma Martínez. Proposal of Orthographic Named Entity Correction in Spanish Voice Queries. *Proceedings of the National Academy of Sciences (PNAS)*, 106(36):15274–15278, 2009. 2.2
- [19] Julián Moreno Schneider. Propuesta de corrección ortográfica de entidades en preguntas por voz en un sistema de QA. Trabajo de final de master, Universidad Carlos III de Madrid, 2011. 2.2
- [20] NIST Information Access Division tools. <http://www.itl.nist.gov/iad/mig//tools>. Último acceso en abril del 2013. 4.2
- [21] Sherpa. <http://sher.pa>. Último acceso en febrero del 2013. 2.1.4
- [22] Siri. <http://www.apple.com/ios/siri>. 2.1.1

- [23] Siri FAQ. <http://www.apple.com/ios/siri/siri-faq>. 2.1.1
- [24] Sun Fire™X4150 Server Service Manual. <http://docs.oracle.com/cd/E19121-01/sf.x4150/820-1852-13>, 2009. Último acceso octubre 2012. 3.2

Apéndice A

Requisitos

A continuación se listan los requisitos tenidos en cuenta a la hora de diseñar e implementar el proyecto.

A.1. Requisitos de Software

A continuación se listan los requisitos que debe cumplir el software. Dichos requisitos se presentan en una tabla con el siguiente formato:

Identificador	Nombre
Fuente	
Prioridad	
Estable	
Descripción	

Cuadro A.1: Definición de la tabla de requisitos de software

donde cada campo significa lo siguiente:

Identificador: código que identifica de forma unívoca cada uno de los requisitos.

Nombre: nombre descriptivo del requisito.

Fuente: origen del requisito.

Prioridad: indica el orden de cumplimiento del requisito. Puede tomar los valores *alta*, *media* y *baja*. A mayor prioridad, mayor es la urgencia para realizarlo.

Estable: indica la probabilidad de cambios en un requisito. Puede tomar los valores *sí* y *no*.

Descripción: explicación del requisito.

ReqSW.1	Sistema operativo
Fuente	Especificaciones de Dragon Naturally Speaking y Windows Speech Recognizer
Prioridad	Alta
Estable	Si
Descripción	Los motores de transcripción incluidos por defecto limitan su soporte a sistemas Windows, sindow Windows Vista y Windows 7 las versiones soportadas por ambos. En el caso de DNS, el soporte se limita a versiones de 32 bits.

Cuadro A.2: Requisito de software 1: sistema operativo

ReqSW.2	Windows Speech Recognizer
Fuente	Servicio web de transcripción
Prioridad	Alta
Estable	Si
Descripción	Uno de los motores ofrecidos por defecto en el servicio web de transcripción es el de Windows Speech Recognizer por lo que será necesario tener instalados los componentes necesarios para su funcionamiento.

Cuadro A.3: Requisito de software 2: Windows Speech Recognizer

A.2. Requisitos de Hardware

En esta sección se listan los requisitos de hardware del proyecto. Para ello se usa una tabla con el mismo formato que [A.1](#)

ReqSW.3	Java Runtime Environment 6
Fuente	Servicio web de transcripción
Prioridad	Alta
Estable	Si
Descripción	El servicio web de transcripción se apoya en algunos programas desarrollados para la versión 6 de Java. Es necesario tener instalado el entorno de ejecución para dicha versión o para una versión compatible.

Cuadro A.4: Requisito de software 3: Java Runtime Environment 6

ReqSW.4	Servidor web
Fuente	Servicio web de transcripción
Prioridad	Alta
Estable	Si
Descripción	Dado que el sistema de transcripción está desarrollado en PHP, para poder ejecutarlo es necesario disponer de un servidor web que sea capaz de interpretar dicho lenguaje.

Cuadro A.5: Requisito de software 4: Servidor web

ReqHW.1	CPU
Fuente	Especificaciones de Windows 7 [11]
Prioridad	Alta
Estable	Si
Descripción	Los requisitos mínimos de hardware para Windows 7 indican que se necesita una CPU de 32 o 64 bit a 1 GHz.

Cuadro A.6: Requisito de hardware 1: CPU

ReqHW.2	RAM
Fuente	Especificaciones de Windows Vista [14] y Windows 7 [11]
Prioridad	Alta
Estable	Si
Descripción	Los requisitos mínimos para poder ejecutar los sistemas operativos soportados por Dragon Naturally Speaking y WSR son de 512 MB en el caso de Windows Vista y 1 GB en el caso de Windows 7.

Cuadro A.7: Requisito de hardware 2: RAM

ReqHW.3	Espacio en disco
Fuente	Especificaciones de Windows Vista [14] y Windows 7 [11]
Prioridad	Alta
Estable	Si
Descripción	En el caso de los sistemas operativos es necesario disponer de un disco duro de 20 GB con al menos 15 GB libres (Windows Vista) o con 16 GB libres (Windows 7).

Cuadro A.8: Requisito de hardware 3: espacio en disco

Apéndice B

Planificación y presupuesto

Esta sección detalla el presupuesto estimado para la ejecución de este proyecto. Primero se discuten las fases y organización del proyecto, y después el coste total del mismo teniendo en cuenta tanto los costes de material como de personal.

B.1. Fases del proyecto

Este proyecto puede ser dividido en tres fases:

Implementación del transcriptor: en esta fase se desarrollan los programas encargados de transcribir las entradas de audio. La gran dificultad de esta fase radica en el desconocimiento de la tecnología empleada (Windows Speech Recognition (WSR)).

Integración con el corrector: durante esta fase se integra el sistema de transcripción con el corrector descrito en [2.2](#) para que la salida del primero sirva como entrada al segundo. Esta fase implica realizar algunos cambios en el sistema de corrección.

Creación de capa web: esta es la última fase del proyecto y en ella se dota de interfaz web a todo el sistema. Esta fase implica la instalación y configuración del servidor web y algunos cambios en los sistemas desarrollados en las dos primeras fases.

Cada fase se puede subdividir en cuatro tareas:

Diseño: en esta tarea se diseñan tanto las interfaces (en caso de que sea necesario) como la estructura interna. En caso de que sea necesario también se lleva a cabo un proceso investigación y aprendizaje para conocer las tecnologías que se van a emplear.

Implementación: se programan los diseños obtenidos en la anterior tarea.

Pruebas: durante esta fase se prueba el software implementado.

Documentación: se recopilan las decisiones tomadas durante cada una de las tareas.

Teniendo en cuenta los conocimientos previos para afrontar cada fase y la complejidad estimada de cada una de ellas, se ha calculado que la primera y la última fase requerirán mayor esfuerzo que la segunda. La figura B.1 muestra el diagrama de Gantt con las fases del proyecto y sus tareas teniendo en cuenta estas consideraciones.

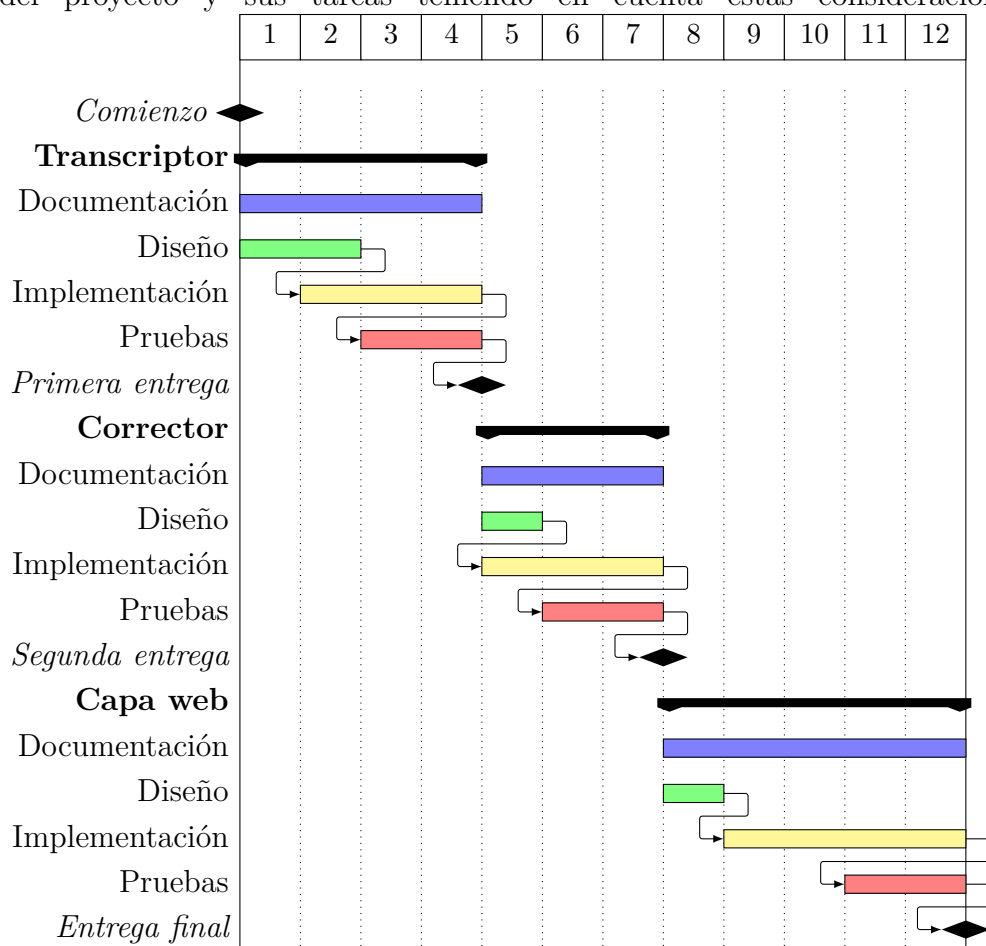


Figura B.1: Planificación del proyecto

B.2. Gastos de materiales

Los gastos de material sólo contemplan el presupuesto destinado a todo lo necesario para desarrollar el proyecto. Eso incluye servidores, programas y consumibles. Dado que el equipamiento usado para el desarrollo del proyecto ya estaba disponible como parte de los recursos pertenecientes al departamento, no se ha realizado ningún desembolso. Aún así se ha creído conveniente incluirlos para tener una idea más completa del coste real. La tabla B.1 refleja el coste estimado en euros.

Cuadro B.1: Material expenses

Concepto	Coste/unidad	Total
Ordenador sobremesa	399,00	399,00
Servidor web	2.555,00	2.555,00
Dragon NS	599,99	599,99
Total		3.553,99

B.3. Gastos de personal

El siguiente coste a analizar es el coste de personal. Para este proyecto se ha contado con un ingeniero que ha trabajado durante las 12 semanas que dura el proyecto. Para la estimación se ha considerado una jornada laboral de 8 horas, una semana de 5 jornadas y un salario medio de 50€ la hora. Sin embargo no todas las tareas requieren el 100 % del tiempo o el esfuerzo y el ingeniero puede llevar a cabo varias tareas a la vez. El coste estimado está recogido en la tabla B.2.

Cuadro B.2: Presupuesto para personal

Concepto	Horas	Esfuerzo	Importe
Diseño (1ª fase)	40	100 %	2.000
Implementación (1ª fase)	90	100 %	4.500
Pruebas (1ª fase)	30	50 %	750
Documentación (1ª fase)	40	30 %	600
<i>continúa en la página siguiente</i>			

Cuadro B.2: Presupuesto para personal (cont.)

Concepto	Horas	Esfuerzo	Importe
Diseño (2ª fase)	12	60 %	360
Implementación (2ª fase)	88	80 %	3.520
Pruebas (2ª fase)	10	30 %	150
Documentación (2ª fase)	10	50 %	250
Diseño (3ª fase)	36	80 %	1.440
Implementación (3ª fase)	104	80 %	4.160
Pruebas (3ª fase)	20	80 %	800
Documentación (3ª fase)	40	80 %	1.600
Total			20.130

B.4. Gastos totales

La tabla B.3 muestra el presupuesto final del proyecto sumando el coste de material y de personal. Las cantidades vienen expresadas en euros.

Cuadro B.3: Presupuesto total

Concepto	Coste
Presupuesto de material	3.553,99
Mano de obra	20.130
Total	23.683.99

Apéndice C

Manual

En este apéndice se detallan las instrucciones de uso de la aplicación así como la información necesaria para llevar a cabo su mantenimiento. Primero se describirán las capas en la que está dividido el servicio y luego se describirá el proceso de instalación y configuración del todo el software necesario para montar el servidor.

C.1. Servicio web

A nivel lógico se puede decir que el proyecto final está dividido en tres capas: un núcleo (el conjunto de aplicaciones que realiza las transcripciones y las correcciones), un API web (para permitir la integración con programas de terceros) y de una interfaz gráfica web. En esta sección se describen esas tres capas sin incluir los ficheros de configuración ya que son descritos en el apartado ??.

C.1.1. Componentes del servicio

A continuación se describen los programas y ficheros que componen el núcleo de la aplicación web. Esto son los programas encargados de realizar la transcripción del audio y de corregir el resultado. También se han incluido en esta sección los programas que aunque no realicen funciones de transcripción o de corrección no proporcionan una interfaz a los usuarios sin acceso directo al servidor.

Transcriptor

Una de las APIs proporcionadas por Microsoft es [SAPI](#) [12]. Este API permite

a los desarrolladores integrar en sus aplicaciones las tecnologías de reconocimiento de voz desarrolladas por Microsoft. Un ejemplo de uso sería Windows Speech Recognition [13] introducido en Windows Vista, aunque en este caso se trate de una aplicación de la misma Microsoft. Una de las funcionalidades proporcionadas es la transcripción de entradas de audio, ya sea directamente del micrófono o a partir de ficheros en formato WAV.

Para poder usar este motor en el servicio se ha desarrollado el comando *transcriptorSAPI* que puede ser invocado de la siguiente forma:

```
transcriptorSAPI -i $input -o $output
```

siendo los parámetros:

\$input (*obligatorio*): ruta a un fichero en formato WAV que contiene el audio que se desea transcribir. Se trata el fichero como si estuviese codificado a 16 kHz, 16 bits, mono.

\$output (*obligatorio*): ruta a un fichero en el que se almacenará el resultado de la transcripción en texto codificado en UTF-8. Este fichero será creado si no existe y en caso contrario será sobrescrito. Está compuesto por las frases reconocidas almacenadas en el siguiente formato:

inicio duración frase

donde *inicio* es el segundo de la locución en el que comienza la frase reconocida, *duración* es la duración de la misma y *frase* es la frase en sí.

Tanto el motor de transcripción de WSR como las interfaces de SAPI permiten la creación y gestión de perfiles de voz así como el entrenamiento de los mismo. Esta opción no ha sido incluida en el comando ya que los intentos implementarlo con gestión de perfiles fueron fallidos.

El flujo de ejecución de comando se puede ver en la figura C.1

Corrección

El sistema de corrección proporcionado ha sido adaptado para recibir como entrada los ficheros devuelto por el sistema de transcripción y devolver como salida el texto transcrito y las posibles alternativas para la corrección. Para facilitar el tratamiento de los ficheros, se ha adaptado la salida del mismo para que el resultado final sea almacenado en un fichero XML con el formato para consumo final por parte del usuario. Dicho formato se puede ver en C.1. Como comando el sistema ha sido agrupado en un fichero jar (llamado

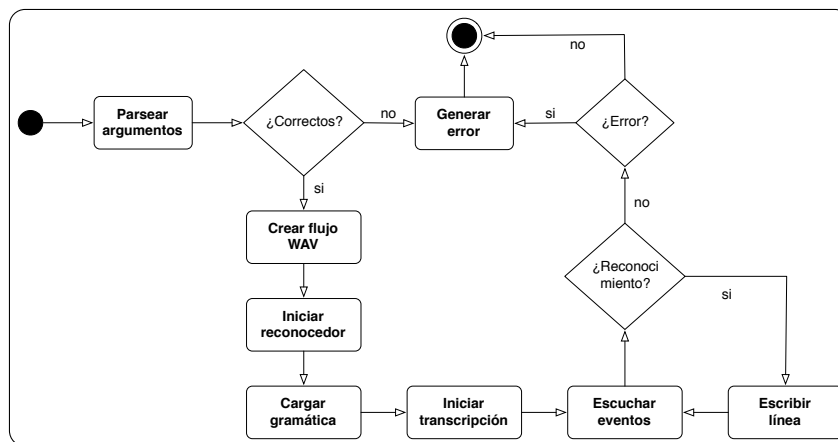


Figura C.1: Diagrama de flujo del transcriptor

GetAlternatives.jar) para facilitar su distribución y ejecución. Toma como parámetros la ruta a un fichero de configuración, la ruta al fichero de resultados devuelto por el sistema de reconocimiento y la ruta al fichero XML de salida.

```
java -jar GetAlternatives.jar conf input output
```

Script intermedio

El fichero batch *transcribe.bat* es necesario para ejecutar en facilmente , de forma secuencial y en background los comandos que llevan a cabo la transcripción y corrección. Los argumentos que necesita son los siguientes:

motor (*obligatorio*): indica cual de los motores de transcripción disponibles se debe usar. Como se ha indicado anteriormente, en este proyecto sólo se hace uso de [SAPI](#) por lo que el valor de este parámetro será siempre *msapi*.

entrada (*obligatorio*): este parámetro se corresponde con la ruta del fichero de entrada excluyendo la extensión del mismo. Dado que sólo trabajamos con ficheros en formato WAV, se asume que esa es la extensión de los ficheros de audio. Los resultados de la transcripción son escritos en ficheros con la misma ruta pero con extensión *txt*, y su corrección se almacenará en ficheros con extensión *xml*.

perfil (*opcional*): indica el perfil de usuario a usar en la transcripción. En este caso no será usado ya que sólo es necesario cuando se utiliza [DNS](#) como motor de transcripción.

Se debe ejecutar con los comandos en el siguiente orden:

```
transcribe.bat $motor $entrada $perfil
```

C.1.2. API del servicio

El siguiente paso en el desarrollo de la plataforma fue dotarla de un [API](#) web para que otras aplicaciones pudiesen hacer uso del servicio. Para ello se ha implementado en el lenguaje PHP una capa capaz de aceptar peticiones HTTP con los parámetros de la transcripción y devolver como resultado un fichero XML que puede ser de error (con el formato que se muestra en [C.1](#)) o contener las correcciones propuestas (con el formato que se muestra en [C.2](#)). Esta capa está implementada en el fichero *transcribe_consulta.php* y hace uso del script descrito en la sección [C.1.1](#).

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
3   <xs:element name="report">
4     <xs:complexType>
5       <xs:all>
6         <xs:element name="code" minOccurs="1" type="xs:int" />
7         <xs:element name="status" minOccurs="1" type="xs:string" />
8       </xs:all>
9     </xs:complexType>
10  </xs:element>
11 </xs:schema>
```

Listing C.1: XDS de los mensajes de error del servicio web

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
3   <xs:element name="transcripciones">
4     <xs:complexType>
5       <xs:sequence>
6         <xs:element name="transcripcion">
7           <xs:complexType>
8             <xs:all>
9               <xs:element name="orden" minOccurs="1" type="xs:int" />
10              <xs:element name="tini" minOccurs="1" type="xs:int" />
11              <xs:element name="tfin" minOccurs="1" type="xs:int" />
12              <xs:element name="persona" minOccurs="1" type="xs:string" />
13              <xs:element name="texto" minOccurs="1" type="xs:string" />
14            </xs:all>
15          </xs:complexType>
16        </xs:element>
17      </xs:sequence>
18    </xs:complexType>
19  </xs:element>
20 </xs:schema>
```

Listing C.2: XDS de los resultados de reconocimiento

Dado que este servicio sólo posee una funcionalidad, el [API](#) se reduce a una llamada a la ruta `transcribe_consulta.php` del servidor donde esté instalado el servicio (e.j.: `http://163.117.129.214/pfc/transcribe_consulta.php`) usando el método POST de HTTP con los siguientes parámetros:

file (*obligatorio*): el fichero WAV que se desea transcribir.

engine (*obligatorio*): indica cual de los motores de transcripción disponibles se debe usar. Los valores posibles en la configuración por defecto son *msapi* para WSR y *dragon* para DNS. Como se ha indicado anteriormente, este trabajo se ha centrado en el uso de WSR, por lo que se asume que el valor de este parámetro es siempre *msapi*.

C.1.3. Interfaz web

Una vez definidas las llamadas disponibles se desarrolla una interfaz gráfica web como ejemplo de aplicación que hace uso del servicio. Esta interfaz (que se muestra en la figura [C.2](#)) permite tanto el envío de ficheros WAV pregrabados como el dictado haciendo uso del micrófono incorporado al ordenador.

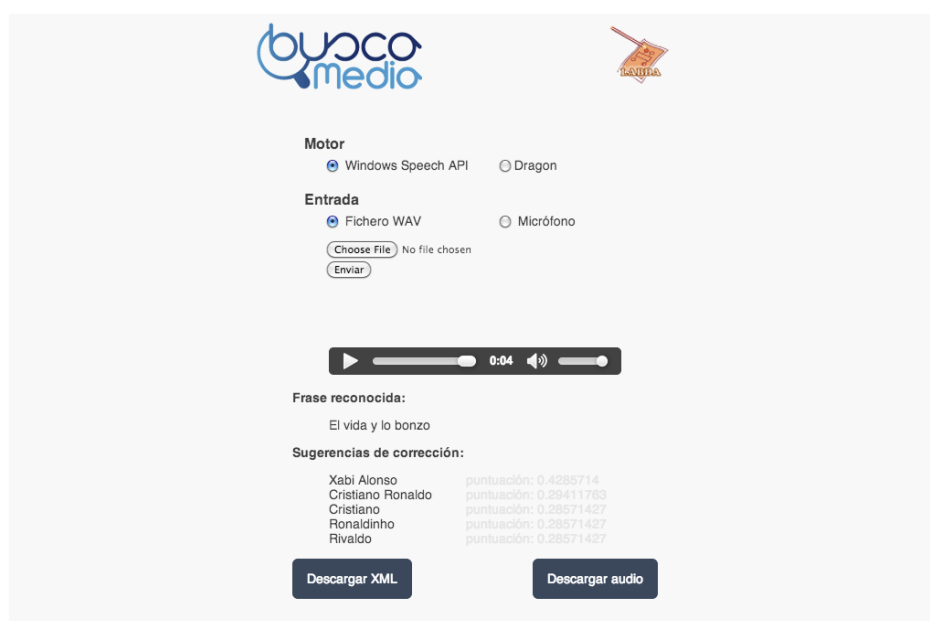


Figura C.2: Interfaz web del sistema

Para las consultas mediante grabaciones se dispone de un botón que también sirve como indicador de estado. Este botón permite arrancar y parar la grabación, y también indicar el estado de la misma. Si el icono mostrado es

el de la figura C.3a entonces se está realizando una grabación. Por el contrario, si el icono mostrado es el de la figura C.3b entonces no hay ninguna grabación en curso.

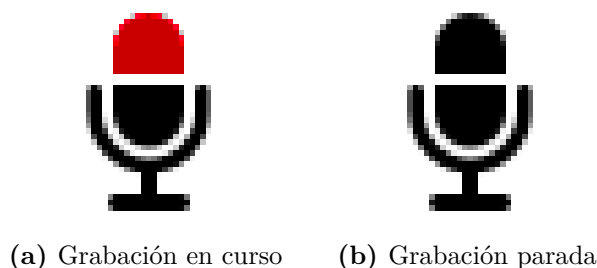


Figura C.3: Iconos de estado de la grabación

Una vez resuelta la consulta, en la misma interfaz aparece un cuadro de texto mostrando el resultado obtenido.

C.2. Instalación de Xen

Una vez desarrollado el software y después de probarlo en la máquina de desarrollo pasamos a instalarlo en un servidor. Dado que no requiere mucho espacio en disco y tampoco mucha memoria RAM no vamos a usar un servidor física dedicado, sino que vamos a aprovechar una disponible. Teniendo en cuenta que el servidor es una maquina de producción y no deseamos interferir con los servicios que proporciona hemos decidido usar una máquina virtual para instalar el nuevo servicio. Dado que el servidor existente tiene instalado un sistema operativo con un kernel Linux nos hemos decantado por usar Xen.

Xen es un monitor de máquinas virtuales. Permite crear y ejecutar varias máquinas virtuales y ejecutar varios sistemas operativos de forma concurrente en la misma máquina física. Esto implica que si una máquina se ve afectada por un problema de software (un virus, un programa que consume excesivos recursos, un bloque...) el resto de las máquinas no se verán afectadas y podrán continuar su ejecución con normalidad.

En esta sección se explica el proceso empleado para instalar Xen. Antes de empezar es necesario tener instalados los siguientes paquetes:

git y mercurial: son herramientas de control de versiones. Son necesarias para poder obtener el código de los parches que se deben aplicar al kernel Linux para usar Xen.

bridge-utils: estas herramientas nos permitirán crear interfaces de red para

las máquinas virtuales y tener acceso a internet con una IP distinta de la de la máquina huesped.

cliente VNC: VNC es un protocolo para el acceso remoto a máquinas. En este caso usaremos un cliente de VNC para conectarnos a las máquinas recién creadas y configurarlas. En este caso se ha optado por `xtightvncviewer`. Si la configuración inicial de la máquina tiene una IP correcta que permita el acceso desde el exterior mediante SSH o RDP este programa no es necesario.

xauth: este paquete permite usar el protocolo de gráficos X sin necesidad de arrancar su servidor. Este protocolo es el más usado para dotar a los sistemas Linux de entorno gráfico y, en nuestro caso, es necesario para poder ejecutar el cliente VNC. Otra opción posible sería instalar y arrancar el servidor X en el sistema anfitrión pero al ser este un servidor y no ser estrictamente necesario el entorno gráfico se ha descartado la opción.

Muchos de los comandos siguientes deben ser ejecutados con permisos de administración por lo que por comodidad se iniciará una sesión de superusuario mediante el comando con el comando `sudo -s`.

En primer lugar es necesario obtener el código fuente parcheado del kernel para poder compilar e instalar el hipervisor. Para ello

1. Nos desplazamos al directorio `/usr/src`, que es el usado habitualmente para almacenar el código fuente de las herramientas que se instalan a partir del código fuente. Para el cambio de directorio ejecutamos el comando `cd /usr/src`
2. Obtenemos el código del parche. Para ello usamos el gestor de versiones `git` para acceder al repositorio del kernel parcheado y clonarlo. En este caso hemos usado el comando `git clone` para guardar en el directorio `linux-2.6-xen` una copia del código alojado en el repositorio `git://git.kernel.org/pub/scm/linux/kernel/git/jeremy/xen.git`
3. Con el comando `cd linux-2.6-xen` nos desplazamos al directorio recién creado y que contiene el código fuente clonado.

Una vez se tenga el repositorio, se debe obtener la versión correspondiente al kernel instalado, que en este caso es de la rama 2.6. Para ello se puede usar el comando `git tag` para ver las versiones estables y el comando `git checkout` para obtener la versión adecuada, que en este caso la 2.6.32.56.

El siguiente paso consiste en compilar el kernel e instalarlo. Para compilarlo se ejecutan la siguiente serie de comandos cuya función es configurar el kernel para nuestra máquina, convertirlo en un paquete e instalarlo:

1. `make menuconfig`
2. `make`
3. `chmod g-s /usr/src -R`
4. `make dep-pkg`
5. `dpkg -i ../linux-image-2.6.32.56_2.6.32.56-2_amd64.deb`
6. `depmod 2.6.32.56`
7. `update-initramfs -c -k 2.6.32.56`

Durante la fase de compilación es posible eliminar los módulos necesarios para controlar cierto hardware si no se encuentra instalado en el servidor (discos SCSI, tabletas, etc). También es posible que se produzca el siguiente error:

```
"dpkg-deb: control directory has bad permissions 2755 (must be >=0755 and <=0775)"
```

Para resolverlo hay que asegurarse de que el directorio en el que se está realizando la compilación tiene los permisos correctos.

Una vez compilado e instalado el kernel es necesario indicarle al sistema que deseamos usarlo para arrancar. Para ello modificamos el fichero `/etc/grub.d/40_custom` para que tenga el siguiente contenido

```

1  #!/bin/sh
2  exec tail -n +3 $0
3  # This file provides an easy way to add custom menu entries.  Simply type the
4  # menu entries you want to add after this comment.  Be careful not to change
5  # the 'exec tail' line above.
6
7  # The dom0_mem directive was added to limit the amount of memory used by Dom0.
8  # This way there is more memory left for the virtual machines.
9
10 menuentry "Xen 4.0 Ubuntu 10.04 Linux 2.6.32.56" {
11     recordfail
12     insmod ext2
13     set root='(hd0,1)'
14     search --no-floppy --fs-uuid --set 3d271eb3-9b34-4dfa-a906-eb805c355db6
15     multiboot (hd0,1)/boot/xen-4.0.gz dummy=dummy dom0_mem=4G
16     module (hd0,1)/boot/vmlinuz-2.6.32.56 dummy=dummy root=UUID=3d271eb3-9b34-4dfa
17     -a906-eb805c355db6 ro quiet
18     module (hd0,1)/boot/initrd.img-2.6.32.56
19 }
```

Listing C.3: Fichero de arranque personalizado

Posteriormente modificamos el fichero `/etc/default/grub` para que el valor de `GRUB_DEFAULT` sea el de la entrada creada en el fichero `40_custom`, incluyendo las comillas, y aplicamos los cambios el comando `update-grub`. En este punto se puede reiniciar la máquina para comprobar que la instalación ha sido correcta.

En el siguiente paso instalaremos las herramientas de Xen para crear y manejar las máquinas virtuales. En este caso también debemos compilarlas así que volvemos al directorio `/usr/src`. Aunque existen varias versiones disponibles en este caso usamos la versión 4.0 porque es la versión estable más reciente. En la página <http://xenbits.xen.org/hg> se pueden ver las versiones estables disponibles. Los siguientes pasos son similares a los usados para obtener el parche de Xen para el kernel, pero en este caso se usa el gestor de versiones Mercurial:

1. Clonamos el repositorio del código usando el comando `hg`:
`hg clone http://xenbits.xensource.com/hg/xen-4.0-testing.hg`
2. Nos cambiamos al directorio con el nuevo repositorio:
`cd xen-4.0-testing.hg`
3. Comprobamos cuales son las versiones estables (releases) disponibles. Podemos usar el comando `hg tags | grep RELEASE` para obtener una lista.
4. Obtenemos la versión estable deseada. En el momento de la realización del proyecto la versión estable era la `RELEASE-4.0.3`
`hg checkout RELEASE-4.0.3`

El siguiente paso consiste en compilar e instalar las herramientas en el sistema usando los siguientes comandos:

1. `make xen && make tools && make stubdom`
2. `make install-xen`
3. `make install-tools PYTHON_PREFIX_ARG=`
4. `make install-stubdom`

Por último, para finalizar con el proceso de instalación debemos modificar los permisos de los scripts que arrancan los dominios de los sistemas operativos huésped y del demonio que los monitoriza. para ellos usamos los comandos `update-rc.d xend defaults 20 21` y `update-rc.d xendomains defaults 2120`. Para comprobar que el proceso se

ha completado correctamente se puede ejecutar el comando `xl list` para listar las máquinas virtuales. Aunque todavía no se haya configurado ni arrancado ninguna máquina virtual el comando debe mostrar la máquina `Domain 0`, que se corresponde con es la máquina física. Una vez configurado el hipervisor, el siguiente paso es configurar las máquinas virtuales. El primer paso consiste en crear un fichero que servirá como disco duro de la máquina. En este caso creamos un fichero de 30 GiB en el directorio `/var/images` usando el comando `dd if=/dev/zero of=/var/images/windows7.img bs=1M count=30720`

Para poder acceder a internet desde las nuevas máquinas, es necesario que el sistema anfitrión se configure para poder incorporar en sobre la marcha varias interfaces de red a una de sus tarjetas. La configuración se las tarjetas de red se encuentra en el fichero `/etc/network/interfaces`. Se debe modificar para cambiar la configuración de la interfaz elegida (en nuestro caso `eth0`) a `allow-hotplug eth0` y reiniciar la conexión de red.

Para definir los parámetros de la máquina virtual debemos crear un fichero de configuración (el usado en este proyecto aparece en el bloque de código [C.4](#)) y colocarlo en el directorio `/etc/xend`. En este punto ya podemos usar la nueva máquina arrancándola con el comando `xl create ruta_del_fichero`. Debemos recordar que en esta primera ejecución la máquina está vacía, por lo que es necesario agregar un dispositivo virtual en el fichero de configuración que contenga la imagen a instalar, como puede ver se en la línea comentada del parámetro `disk` en la configuración de la máquina. Dado que en el proceso de instalación no está configurada la red del sistema huésped deberemos acceder a el desde el sistema anfitrión. Para ello basta conectarse el servidor físico usando el protocolo `ssh` habilitando la retransmisión de gráficos con el protocolo `X` y arrancar un visor `VNC` a la dirección y puerto indicados por los parámetros `vnclisten` y `vncdisplay` respectivamente.

Tras terminar la instalación de la nueva máquina es recomendable cambiar el orden de arranque o incluso eliminar del fichero de configuración el dispositivo virtual con al imagen de instalación. También es recomendable habilitar servidores que nos permitan acceder remotamente y de forma directa a la nueva máquina para administrarla. En este caso se ha habilitado el protocolo `RDP`.

```
1 # -*- mode: python; -*-
2 # Kernel image file.
3 kernel = "hvmloader"
4
5 # The domain build function. HVM domain uses 'hvm'.
6 builder='hvm'
7
8 # Initial memory allocation (in megabytes) for the new domain.
9 memory = 1024
10
11 # Shadow pagetable memory for the domain, in MB.
```

```

12 shadow_memory = 8
13
14 # A name for your domain. All domains must have different names.
15 name = "PUT_NAME_HERE"
16
17 # The number of cpus guest platform has, default=1
18 vcpus=1
19
20 # Optionally define mac and/or bridge for the network interfaces.
21 vif = [ 'type=ioemu, bridge=eth0, ip=PUT_IP_ADDRESS_HERE' ]
22
23 # Define the disk devices you want the domain to have access to, and
24 # what you want them accessible as.
25 disk = [ 'file:/var/images/buscamedia_kojiro.img,hda,w',
26 #       'file:/var/images/operating_system.iso,hdb:cdrom,r',
27         ]
28
29 # Configure the behaviour when a domain exits. There are three 'reasons'
30 # for a domain to stop: poweroff, reboot, and crash.
31 on_poweroff = 'destroy'
32 on_reboot   = 'restart'
33 on_crash    = 'restart'
34
35 # Device Model to be used
36 device_model = 'qemu-dm'
37
38 # boot on floppy (a), hard disk (c), Network (n) or CD-ROM (d)
39 boot="dc"
40
41 # enable SDL library for graphics, default = 0
42 sdl=0
43
44 # enable OpenGL for texture rendering inside the SDL window, default = 1
45 # valid only if sdl is enabled.
46 opengl=1
47
48 # enable VNC library for graphics, default = 1
49 vnc=1
50
51 # address that should be listened on for the VNC server if vnc is set.
52 # default is to use 'vnc-listen' setting from
53 # auxbin.xen_configdir() + /xend-config.sxp
54 vnclisten="127.0.0.1"
55
56 # set password for domain's VNC console
57 # default is depends on vncpasswd in xend-config.sxp
58 vncpasswd=''
59
60 # enable stdvga, default = 0 (use cirrus logic device model)
61 stdvga=1
62 # Use more RAM for video. This needs stdvga=1
63 videoram=16
64
65 # serial port re-direct to pty device, /dev/pts/n
66 # then xm console or minicom can connect
67 serial='pty'
68
69 # tsc_mode : TSC mode (0=default, 1=native TSC, 2=never emulate, 3=pvrdtscp)
70 tsc_mode=0
71
72 # Set keyboard layout, default is en-us keyboard.
73 keymap='es'

```

Listing C.4: Fichero de configuración para la máquina virtual

C.3. Servidor web

Una vez que se las máquinas virtuales han sido instalada y configuradas, el siguiente paso consiste en instalar y configurar el servidor web. Como se ha mencionado en el apartado 3.2, se ha optado por usar el paquete WAMP, que proporciona un servidor HTTP Apache y un intérprete de PHP, pero se puede usar cualquier otro servidor web capaz de interpretar ficheros PHP.

La instalación de la aplicación es sencilla. Basta con almacenar los ficheros que la componen en una ruta pública del servidor para que pueda empezar a usarse, pero para que funcione correctamente es posible que sea necesario realizar algunas modificaciones en el fichero de configuración de PHP.

El primer paso para realizar las modificaciones es localizar el fichero que se está usado como fichero de configuración. Para ello se puede crear un fichero PHP en el servidor que contenga la siguiente llamada:

```
<?php phpinfo(); ?>
```

Al acceder al fichero a través del servidor HTTP se mostrará una tabla similar a la que se puede ver en la figura C.4 y que contiene información sobre la configuración del módulo de PHP para el servidor. En este caso se puede ver que el fichero de configuración (general mente llamado *php.ini*) está guardado en el directorio */etc*. Una vez localizado el fichero se debe comprobar que las siguientes opciones están configuradas de seguir los siguientes criterios:

fileinfo

Esta extensión del intérprete de PHP permite obtener información sobre los ficheros (tamaño, tipo...) y suele estar desactiva por defecto. Es necesario tenerla activada ya que el servicio hace uso de esta funcionalidad.

memory_limit

Este parámetro indica la cantidad de memoria dedicada a cada proceso. Dado que estamos tratando con ficheros es recomendable que su valor sea como mínimo 128M.

post_max_size

Con este parámetro se limita el tamaño dedicado al contenido de las peticiones POST. Por lo general el tamaño de estas peticiones no suele ser

grande, pero al estar tratando con fichero se incrementa notablemente. Se ha usado 100M, que es un valor lo suficientemente grande como para disminuir la probabilidad de que se produzcan fallo por peticiones grandes.

upload_max_filesize

Este parámetro controla el tamaño máximo de los ficheros enviados al servidor a través del protocolo HTTP que acepta el intérprete de PHP. En principio los ficheros de audio manejados no deberían pasar de los 30MiB, pero es posible que un usuario envíe por error un fichero más grande. Para reducir las posibilidades de error se ha usado 80M como valor del parámetro.

Se debe tener en cuenta que la final se debe cumplir

$$memory_limit > post_max_size > upload_max_size$$


<div> <div>PHP Version 5.3.15</div>  </div>	
System	Darwin epona.local 10.8.0 Darwin Kernel Version 10.8.0: Tue Jun 7 16:33:36 PDT 2011; root:xnu-1504.15.3~1/RELEASE_I386 i386
Build Date	Jul 31 2012 18:37:44
Configure Command	'/private/var/tmp/apache_mod_php/apache_mod_php-53.10~1/php/configure' '--prefix=/usr' '--mandir=/usr/share/man' '--infodir=/usr/share/info' '--disable-dependency-tracking' '--sysconfdir=/private/etc' '--with-apxs2=/usr/sbin/apxs' '--enable-cli' '--with-config-file-path=/etc' '--with-libxml-dir=/usr' '--with-openssl=/usr' '--with-kerberos=/usr' '--with-zlib=/usr' '--enable-bcmath' '--with-bz2=/usr' '--enable-calendar' '--with-curl=/usr' '--enable-exif' '--enable-ftp' '--with-gd' '--with-jpeg-dir=/BinaryCache/apache_mod_php/apache_mod_php-53.10~1/Root/usr/local' '--with-png-dir=/BinaryCache/apache_mod_php/apache_mod_php-53.10~1/Root/usr/local' '--enable-gd-native-ttf' '--with-ldap=/usr' '--with-ldap-sasl=/usr' '--enable-mbstring' '--enable-mbregex' '--with-mysql=mysqlnd' '--with-mysqli=mysqlnd' '--with-pdo-mysql=mysqlnd' '--with-mysql-sock=/var/mysql/mysql.sock' '--without-pear' '--with-iodbc=/usr' '--enable-shmop' '--with-snmp=/usr' '--enable-soap' '--enable-sockets' '--enable-sysvmsg' '--enable-sysvsem' '--enable-sysvshm' '--enable-wddx' '--with-xmlrpc' '--with-iconv-dir=/usr' '--with-xsl=/usr' '--enable-zend-multibyte' '--enable-zip' '--with-pcre-regex=/usr'
Server API	Apache 2.0 Handler
Virtual Directory Support	disabled
Configuration File (php.ini) Path	/etc
Loaded Configuration File	(none)
Scan this dir for additional .ini files	(none)
Additional .ini files parsed	(none)

Figura C.4: Fragmento de la salida de phpinfo

```
www/
├── acceptFile.php
├── buscamedia.inc
├── commands/
│   ├── GetAlternatives.jar
│   ├── java/
│   │   └── fileconverter/
│   │       └── SAPItoXMLConverter.class
│   └── transcriptorSAPI.exe
├── images/
│   ├── buscamedia.png
│   ├── iconRecord.png
│   ├── iconStop.png
│   └── labda.png
├── index.php
├── js/
│   ├── JRecorder.swf
│   ├── jRecorder.js
│   ├── jquery-1.7.2.js
│   ├── jquery-ui-1.8.20.custom.min.js
│   └── script.js
├── resources/
│   ├── buscavoice.properties
│   └── databasefile.txt
├── style/
│   └── buscamedia.css
├── transcribe.bat
└── transcribe_consulta.php
```

Figura C.5: Contenido del directorio del servicio

Apéndice D

Posibles mejoras

Durante el desarrollo de la práctica han surgido varias ideas para mejorar el servicio pero no han sido implementadas. Dichas ideas se recogen en esta sección.

Uso de perfiles

Según la documentación sobre [WSR](#) disponible, es posible emplear perfiles de aprendizaje para entrenar el reconocedor y mejorar la tasa de aciertos. Se ha intentado incluir dicha funcionalidad en el comando de transcripción desarrollado pero no ha sido posible por falta de conocimientos. Consideramos que hacer uso de dichos perfiles mejoraría bastante los resultados obtenidos.

Estandarizar los comandos de transcripción

A lo largo de este documento se ha estado hablando de un solo comando para las transcripciones, pero es posible que en un futuro se desean agregar más motores y eso implicaría desarrollar más comandos. Creemos que sería recomendable que los comandos tuviesen la misma sintaxis para los parámetros y que como mínimo implementen estos:

- Perfil a usar [opcional]. En caso de que se desee indicar un perfil para realizar la transcripción se usará la etiqueta **-p perfil** para indicarlo. Este parámetro es opcional porque no está garantizado que todos los motores que se puedan usar tengan soporte para perfiles y porque consideramos que sería recomendable que en caso de permitir el uso de perfiles se tuviese un perfil por defecto.

- Ruta al fichero wav de entrada [obligatorio]. Este parámetro es necesario porque indica la ruta al fichero que se desea transcribir. Para indicarlo se usará la etiqueta `-i ruta`.
- Ruta al fichero de salida [opcional]. Este parámetro indicará el fichero donde escribir el resultado de la transcripción. Es opcional porque es posible que el usuario necesite redirigir la salida a otro comando para lo cual es recomendable usar la salida estándar. La etiqueta para este parámetro será `-o ruta`.

Cada comando podrá implementar otros parámetros en caso de que los necesite pero deberán ser opcionales.

Registro de más comandos

El mecanismo actual para agregar más comandos a la aplicación es redundante. Se debe modificar la interfaz web de la aplicación, el fichero con el [API](#) y el fichero bat. Es recomendable simplificar el mecanismo para registrar más comandos para reducir la posibilidad de cometer errores y facilitar el mantenimiento. Dado que la aplicación está pensada para la web proponemos crear el fichero `commands.php` con un array de comandos con el siguiente formato:

```
1 <?php
2 $commands= [
3     [
4         "path":"ruta al fichero ejecutable",
5         "profiles": ["perfil1", "perfil2",...,"perfilN"]
6         // si sólo hay un perfil se puede omitir la lista de perfiles
7     ],
8     [...]
9 ];
10 ?>
```

Listing D.1: Propuesta de formato para el registro de comandos de transcripción

Bases de datos

En la actualidad la aplicación no hace uso de ningún sistema para gestionar la persistencia de los datos. El resultado de las transcripciones de almacena en ficheros XML en el sistema de ficheros de Windows y los registros con los valores para las correcciones se almacenan en ficheros de texto plano que no están optimizados para accesos aleatorios. Consideramos que incluir una base de datos en la aplicación podría mejorarla ya que:

- Se pueden almacenar los datos del sistema de corrección en una base de datos. Eso reduciría los tiempo de acceso y permitiría usar los mismos datos de corrección en varias máquinas distintas usando una de ellas como servidor de datos.
- Se pueden almacenar los resultados de la transcripción en la base de datos y evitar posibles problemas de permisos de acceso por modificaciones en el sistema de ficheros.
- Se podrían almacenar datos extra como el tiempo empleado en resolver cada consulta, informes de errores, etc.

REST

En el desarrollo actual no se ha usado ninguna arquitectura concreta a la hora de desarrollar el API del servicio web. Se podría volver a implementar el servicio haciendo uso de REST, que es un estilo de arquitectura para aplicaciones web sencillo y bastante usado hoy en día. Consiste en usar las URLs del servicio junto con los métodos y códigos de error HTTP para determinar la funcionalidad. Eso implica una configuración determinada en el lado del servidor.

También se podría usar SOAP pero no lo recomendamos porque el servicio es muy sencillo y eso implicaría sobrecargas y complicaciones innecesarias.

Aprendizaje

Se podría agregar un sistema básico para recibir retroalimentación de los usuarios e incorporar un sistema de aprendizaje en la aplicación. Los usuarios podrían indicar cual de las opciones de corrección propuestas es la correcta y de esa forma el sistema se iría ajustando automáticamente.

Apéndice E

Salida de las pruebas

En este anexo se listan los ficheros de resultados devueltos por la aplicación al realizar las pruebas.

E.1. Resultados del servicio web

```
1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="58" end="219" person="UNKNOWN">
4     <correction>
5       <original>Andrés iniesta</original>
6       <alternative score="0.6666666">Andrés Iniesta</alternative>
7       <alternative score="0.5">Iniesta</alternative>
8       <alternative score="0.25">Casillas</alternative>
9       <alternative score="0.25">Benzema</alternative>
10      <alternative score="0.25">Canales</alternative>
11      <alternative score="0.25">Dani Alves</alternative>
12      <alternative score="0.25">Rosenberg</alternative>
13      <alternative score="0.25">Guardiola</alternative>
14      <alternative score="0.25">Keita</alternative>
15      <alternative score="0.25">Pedro León</alternative>
16      <alternative score="0.25">Sneijder</alternative>
17      <alternative score="0.25">Llorente</alternative>
18      <alternative score="0.25">Pinto</alternative>
19      <alternative score="0.25">Dani Güiza</alternative>
20      <alternative score="0.25">Xavi Prieto</alternative>
21      <alternative score="0.25">Lass Diarra</alternative>
22      <alternative score="0.23076922">Andrés Palop</alternative>
23      <alternative score="0.21428573">Van Nistelrooy</alternative>
24      <alternative score="0.19999999">Rafael Van der Vaart</alternative>
25      <alternative score="0.19999999">Fernando Torres</alternative>
26      <alternative score="0.17647058">Cristiano Ronaldo</alternative>
27      <alternative score="0.16666669">Cristiano</alternative>
28      <alternative score="0.16666669">Nolito</alternative>
29      <alternative score="0.16666669">Ronaldinho</alternative>
30      <alternative score="0.16666669">Villa</alternative>
31      <alternative score="0.16666669">David Villa</alternative>
32      <alternative score="0.16666669">Messi</alternative>
33      <alternative score="0.16666669">Marcelo</alternative>
34      <alternative score="0.16666669">Leo Messi</alternative>
```

```

35 <alternative score="0.16666669">Lionel Messi</alternative>
36 <alternative score="0.16666669">Alves</alternative>
37 <alternative score="0.16666669">Alcántara</alternative>
38 <alternative score="0.16666669">Santín</alternative>
39 <alternative score="0.16666669">Aduriz</alternative>
40 <alternative score="0.16666669">Valdés</alternative>
41 <alternative score="0.16666669">Cesc</alternative>
42 <alternative score="0.16666669">Aguirreche</alternative>
43 <alternative score="0.16666669">Riki</alternative>
44 <alternative score="0.16666669">Aranzubía</alternative>
45 <alternative score="0.16666669">Griezmann</alternative>
46 <alternative score="0.16666669">Güiza</alternative>
47 <alternative score="0.16666669">David De Gea</alternative>
48 <alternative score="0.16666669">De Gea</alternative>
49 <alternative score="0.15384614">Gerard Piqué</alternative>
50 <alternative score="0.15384614">Iker Casillas</alternative>
51 <alternative score="0.15384614">Didier Drogba</alternative>
52 <alternative score="0.13333333">Xavi Hernández</alternative>
53 <alternative score="0.08333331">Tiago</alternative>
54 <alternative score="0.08333331">Jeffren</alternative>
55 <alternative score="0.08333331">Deco</alternative>
56 <alternative score="0.08333331">Forlán</alternative>
57 <alternative score="0.08333331">Xavi</alternative>
58 <alternative score="0.08333331">Edú</alternative>
59 <alternative score="0.08333331">Müller</alternative>
60 <alternative score="0.08333331">Kun Agüero</alternative>
61 <alternative score="0.08333331">Sergio Ramos</alternative>
62 <alternative score="0.08333331">Carvalho</alternative>
63 <alternative score="0.08333331">Xabi Alonso</alternative>
64 <alternative score="0.08333331">Özil</alternative>
65 <alternative score="0.08333331">Maxwell</alternative>
66 <alternative score="0.08333331">Rivaldo</alternative>
67 <alternative score="0.08333331">César</alternative>
68 <alternative score="0.08333331">Di María</alternative>
69 <alternative score="0.08333331">Pepe</alternative>
70 <alternative score="0.08333331">Carles Puyol</alternative>
71 <alternative score="0.0714286">Miguel García</alternative>
72 <alternative score="0.0">Eto'o</alternative>
73 <alternative score="0.0">Puyol</alternative>
74 <alternative score="0.0">Raúl</alternative>
75 <alternative score="0.0">Palop</alternative>
76 <alternative score="0.0">Soldado</alternative>
77 <alternative score="0.0">Miguel Albiol</alternative>
78 <alternative score="0.0">Diego Forlán</alternative>
79 <alternative score="0.0">Falcao</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.1: Salida para Andrés Iniesta (voz femenina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="56" end="200" person="UNKNOWN">
4     <correction>
5       <original>Cristianos Donaldo</original>
6       <alternative score="0.8235294">Cristiano Ronaldo</alternative>
7       <alternative score="0.4705882">Cristiano</alternative>
8       <alternative score="0.29411763">Rivaldo</alternative>
9       <alternative score="0.29411763">Soldado</alternative>
10      <alternative score="0.29411763">Griezmann</alternative>

```

```

11 <alternative score="0.2352941">Ronaldinho</alternative>
12 <alternative score="0.2352941">Carbalho</alternative>
13 <alternative score="0.2352941">Xabi Alonso</alternative>
14 <alternative score="0.2352941">Canales</alternative>
15 <alternative score="0.2352941">Aranzubía</alternative>
16 <alternative score="0.17647058">Casillas</alternative>
17 <alternative score="0.17647058">Tiago</alternative>
18 <alternative score="0.17647058">Eto'o</alternative>
19 <alternative score="0.17647058">Iniesta</alternative>
20 <alternative score="0.17647058">Iker Casillas</alternative>
21 <alternative score="0.17647058">David Villa</alternative>
22 <alternative score="0.17647058">Andrés Palop</alternative>
23 <alternative score="0.17647058">Sergio Ramos</alternative>
24 <alternative score="0.17647058">Marcelo</alternative>
25 <alternative score="0.17647058">Santín</alternative>
26 <alternative score="0.17647058">Di María</alternative>
27 <alternative score="0.17647058">Rosenberg</alternative>
28 <alternative score="0.17647058">Guardiola</alternative>
29 <alternative score="0.17647058">Keita</alternative>
30 <alternative score="0.17647058">Diego Forlán</alternative>
31 <alternative score="0.17647058">Pinto</alternative>
32 <alternative score="0.17647058">Fernando Torres</alternative>
33 <alternative score="0.17647058">Didier Drogba</alternative>
34 <alternative score="0.17647058">Falcao</alternative>
35 <alternative score="0.17647058">Lass Diarra</alternative>
36 <alternative score="0.14999998">Rafael Van der Vaart</alternative>
37 <alternative score="0.11764705">Nolito</alternative>
38 <alternative score="0.11764705">Villa</alternative>
39 <alternative score="0.11764705">Puyol</alternative>
40 <alternative score="0.11764705">Messi</alternative>
41 <alternative score="0.11764705">Raúl</alternative>
42 <alternative score="0.11764705">Palop</alternative>
43 <alternative score="0.11764705">Kun Agüero</alternative>
44 <alternative score="0.11764705">Benzema</alternative>
45 <alternative score="0.11764705">Özil</alternative>
46 <alternative score="0.11764705">Lionel Messi</alternative>
47 <alternative score="0.11764705">Maxwell</alternative>
48 <alternative score="0.11764705">Alcántara</alternative>
49 <alternative score="0.11764705">César</alternative>
50 <alternative score="0.11764705">Valdés</alternative>
51 <alternative score="0.11764705">Dani Alves</alternative>
52 <alternative score="0.11764705">Miguel Albiol</alternative>
53 <alternative score="0.11764705">Xavi Hernández</alternative>
54 <alternative score="0.11764705">Sneijder</alternative>
55 <alternative score="0.11764705">Riki</alternative>
56 <alternative score="0.11764705">Llorente</alternative>
57 <alternative score="0.11764705">Dani Güiza</alternative>
58 <alternative score="0.11764705">Güiza</alternative>
59 <alternative score="0.11764705">David De Gea</alternative>
60 <alternative score="0.11764705">Xavi Prieto</alternative>
61 <alternative score="0.11764705">Carles Puyol</alternative>
62 <alternative score="0.11764705">Van Nistelrooy</alternative>
63 <alternative score="0.058823526">Miguel García</alternative>
64 <alternative score="0.058823526">Jeffren</alternative>
65 <alternative score="0.058823526">Deco</alternative>
66 <alternative score="0.058823526">Gerard Piqué</alternative>
67 <alternative score="0.058823526">Forlán</alternative>
68 <alternative score="0.058823526">Xavi</alternative>
69 <alternative score="0.058823526">Edú</alternative>
70 <alternative score="0.058823526">Müller</alternative>
71 <alternative score="0.058823526">Andrés Iniesta</alternative>
72 <alternative score="0.058823526">Leo Messi</alternative>

```

```

73 <alternative score="0.058823526">Alves</alternative>
74 <alternative score="0.058823526">Aduriz</alternative>
75 <alternative score="0.058823526">Pedro León</alternative>
76 <alternative score="0.058823526">Cesc</alternative>
77 <alternative score="0.058823526">Aguirreche</alternative>
78 <alternative score="0.058823526">De Gea</alternative>
79 <alternative score="0.0">Pepe</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.2: Salida para Cristiano Ronalvo (voz femenina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="56" end="176" person="UNKNOWN">
4     <correction>
5       <original>David villa</original>
6       <alternative score="0.6363636">David Villa</alternative>
7       <alternative score="0.39999998">Casillas</alternative>
8       <alternative score="0.39999998">Villa</alternative>
9       <alternative score="0.39999998">Guardiola</alternative>
10      <alternative score="0.3333333">David De Gea</alternative>
11      <alternative score="0.3076923">Iker Casillas</alternative>
12      <alternative score="0.3">Carbalho</alternative>
13      <alternative score="0.3">Maxwell</alternative>
14      <alternative score="0.27272725">Xabi Alonso</alternative>
15      <alternative score="0.27272725">Dani Güiza</alternative>
16      <alternative score="0.27272725">Xavi Prieto</alternative>
17      <alternative score="0.27272725">Lass Diarra</alternative>
18      <alternative score="0.21428573">Van Nistelrooy</alternative>
19      <alternative score="0.19999999">Cristiano</alternative>
20      <alternative score="0.19999999">Xavi</alternative>
21      <alternative score="0.19999999">Iniesta</alternative>
22      <alternative score="0.19999999">Raúl</alternative>
23      <alternative score="0.19999999">Palop</alternative>
24      <alternative score="0.19999999">Marcelo</alternative>
25      <alternative score="0.19999999">Özil</alternative>
26      <alternative score="0.19999999">Andrés Iniesta</alternative>
27      <alternative score="0.19999999">Rivaldo</alternative>
28      <alternative score="0.19999999">Valdés</alternative>
29      <alternative score="0.19999999">Canales</alternative>
30      <alternative score="0.19999999">Dani Alves</alternative>
31      <alternative score="0.19999999">Keita</alternative>
32      <alternative score="0.19999999">Riki</alternative>
33      <alternative score="0.19999999">Aranzubía</alternative>
34      <alternative score="0.19999999">Güiza</alternative>
35      <alternative score="0.19999999">Falcao</alternative>
36      <alternative score="0.15384614">Gerard Piqué</alternative>
37      <alternative score="0.15384614">Andrés Palop</alternative>
38      <alternative score="0.15384614">Miguel Albiol</alternative>
39      <alternative score="0.15384614">Didier Drogba</alternative>
40      <alternative score="0.14999998">Rafael Van der Vaart</alternative>
41      <alternative score="0.13333333">Xavi Hernández</alternative>
42      <alternative score="0.11764705">Cristiano Ronaldo</alternative>
43      <alternative score="0.10000024">Nolito</alternative>
44      <alternative score="0.10000024">Tiago</alternative>
45      <alternative score="0.10000024">Ronaldinho</alternative>
46      <alternative score="0.10000024">Puyol</alternative>
47      <alternative score="0.10000024">Edú</alternative>
48      <alternative score="0.10000024">Messi</alternative>

```



```

49 <alternative score="0.10000024">Müller</alternative>
50 <alternative score="0.10000024">Benzema</alternative>
51 <alternative score="0.10000024">Alcántara</alternative>
52 <alternative score="0.10000024">Santín</alternative>
53 <alternative score="0.10000024">Soldado</alternative>
54 <alternative score="0.10000024">Aduriz</alternative>
55 <alternative score="0.10000024">Di María</alternative>
56 <alternative score="0.10000024">Rosenberg</alternative>
57 <alternative score="0.10000024">Sneijder</alternative>
58 <alternative score="0.10000024">Aguirreche</alternative>
59 <alternative score="0.10000024">Griezmann</alternative>
60 <alternative score="0.10000024">Pinto</alternative>
61 <alternative score="0.10000024">De Gea</alternative>
62 <alternative score="0.08333331">Sergio Ramos</alternative>
63 <alternative score="0.08333331">Carles Puyol</alternative>
64 <alternative score="0.07692307">Diego Forlán</alternative>
65 <alternative score="0.0714286">Miguel García</alternative>
66 <alternative score="0.06666666">Fernando Torres</alternative>
67 <alternative score="0.0">Jeffren</alternative>
68 <alternative score="0.0">Eto'o</alternative>
69 <alternative score="0.0">Deco</alternative>
70 <alternative score="0.0">Forlán</alternative>
71 <alternative score="0.0">Kun Agüero</alternative>
72 <alternative score="0.0">Leo Messi</alternative>
73 <alternative score="0.0">Lionel Messi</alternative>
74 <alternative score="0.0">Alves</alternative>
75 <alternative score="0.0">César</alternative>
76 <alternative score="0.0">Pepe</alternative>
77 <alternative score="0.0">Pedro León</alternative>
78 <alternative score="0.0">Cesc</alternative>
79 <alternative score="0.0">Llorente</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.3: Salida para David Villa (voz femenina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="55" end="196" person="UNKNOWN">
4     <correction>
5       <original>Iker casillas</original>
6       <alternative score="0.7692308">Iker Casillas</alternative>
7       <alternative score="0.5833334">Casillas</alternative>
8       <alternative score="0.3333333">Villa</alternative>
9       <alternative score="0.3333333">David Villa</alternative>
10      <alternative score="0.3333333">Guardiola</alternative>
11      <alternative score="0.25">Messi</alternative>
12      <alternative score="0.25">Carvalho</alternative>
13      <alternative score="0.25">Canales</alternative>
14      <alternative score="0.25">Keita</alternative>
15      <alternative score="0.25">Falcao</alternative>
16      <alternative score="0.23076922">Andrés Palop</alternative>
17      <alternative score="0.23076922">Didier Drogba</alternative>
18      <alternative score="0.21428573">Miguel García</alternative>
19      <alternative score="0.19999999">Andrés Iniesta</alternative>
20      <alternative score="0.19999999">Fernando Torres</alternative>
21      <alternative score="0.16666669">Cristiano</alternative>
22      <alternative score="0.16666669">Forlán</alternative>
23      <alternative score="0.16666669">Xavi</alternative>
24      <alternative score="0.16666669">Iniesta</alternative>

```

```

25 <alternative score="0.16666669">Raúl</alternative>
26 <alternative score="0.16666669">Müller</alternative>
27 <alternative score="0.16666669">Palop</alternative>
28 <alternative score="0.16666669">Sergio Ramos</alternative>
29 <alternative score="0.16666669">Benzema</alternative>
30 <alternative score="0.16666669">Marcelo</alternative>
31 <alternative score="0.16666669">Özil</alternative>
32 <alternative score="0.16666669">Maxwell</alternative>
33 <alternative score="0.16666669">Alves</alternative>
34 <alternative score="0.16666669">Rivaldo</alternative>
35 <alternative score="0.16666669">César</alternative>
36 <alternative score="0.16666669">Aduriz</alternative>
37 <alternative score="0.16666669">Valdés</alternative>
38 <alternative score="0.16666669">Di María</alternative>
39 <alternative score="0.16666669">Dani Alves</alternative>
40 <alternative score="0.16666669">Sneijder</alternative>
41 <alternative score="0.16666669">Cesc</alternative>
42 <alternative score="0.16666669">Riki</alternative>
43 <alternative score="0.16666669">Güiza</alternative>
44 <alternative score="0.16666669">De Gea</alternative>
45 <alternative score="0.15384614">Gerard Piqué</alternative>
46 <alternative score="0.15384614">Miguel Albiol</alternative>
47 <alternative score="0.15384614">Diego Forlán</alternative>
48 <alternative score="0.14999998">Rafael Van der Vaart</alternative>
49 <alternative score="0.14285713">Van Nistelrooy</alternative>
50 <alternative score="0.13333333">Xavi Hernández</alternative>
51 <alternative score="0.11764705">Cristiano Ronaldo</alternative>
52 <alternative score="0.08333331">Nolito</alternative>
53 <alternative score="0.08333331">Tiago</alternative>
54 <alternative score="0.08333331">Jeffren</alternative>
55 <alternative score="0.08333331">Ronaldinho</alternative>
56 <alternative score="0.08333331">Deco</alternative>
57 <alternative score="0.08333331">Puyol</alternative>
58 <alternative score="0.08333331">Xabi Alonso</alternative>
59 <alternative score="0.08333331">Leo Messi</alternative>
60 <alternative score="0.08333331">Alcántara</alternative>
61 <alternative score="0.08333331">Santín</alternative>
62 <alternative score="0.08333331">Soldado</alternative>
63 <alternative score="0.08333331">Pepe</alternative>
64 <alternative score="0.08333331">Pedro León</alternative>
65 <alternative score="0.08333331">Llorente</alternative>
66 <alternative score="0.08333331">Aranzubía</alternative>
67 <alternative score="0.08333331">Griezmann</alternative>
68 <alternative score="0.08333331">Pinto</alternative>
69 <alternative score="0.08333331">Dani Güiza</alternative>
70 <alternative score="0.08333331">Xavi Prieto</alternative>
71 <alternative score="0.08333331">Carles Puyol</alternative>
72 <alternative score="0.08333331">Lass Diarra</alternative>
73 <alternative score="0.0">Eto'o</alternative>
74 <alternative score="0.0">Edú</alternative>
75 <alternative score="0.0">Kun Agüero</alternative>
76 <alternative score="0.0">Lionel Messi</alternative>
77 <alternative score="0.0">Rosenberg</alternative>
78 <alternative score="0.0">Aguirreche</alternative>
79 <alternative score="0.0">David De Gea</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.4: Salida para Iker Casillas (voz femenina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="52" end="194" person="UNKNOWN">
4     <correction>
5       <original>Lars diana</original>
6       <alternative score="0.45454544">Lass Diarra</alternative>
7       <alternative score="0.44444442">Cristiano</alternative>
8       <alternative score="0.44444442">Guardiola</alternative>
9       <alternative score="0.36363637">David Villa</alternative>
10      <alternative score="0.33333333">Casillas</alternative>
11      <alternative score="0.3">Alcántara</alternative>
12      <alternative score="0.29411763">Cristiano Ronaldo</alternative>
13      <alternative score="0.26666665">Andrés Iniesta</alternative>
14      <alternative score="0.25">David De Gea</alternative>
15      <alternative score="0.25">Carles Puyol</alternative>
16      <alternative score="0.23076922">Gerard Piqué</alternative>
17      <alternative score="0.23076922">Iker Casillas</alternative>
18      <alternative score="0.23076922">Andrés Palop</alternative>
19      <alternative score="0.22222221">Tiago</alternative>
20      <alternative score="0.22222221">Forlán</alternative>
21      <alternative score="0.22222221">Xavi</alternative>
22      <alternative score="0.22222221">Villa</alternative>
23      <alternative score="0.22222221">Messi</alternative>
24      <alternative score="0.22222221">Marcelo</alternative>
25      <alternative score="0.22222221">Carbalho</alternative>
26      <alternative score="0.22222221">Santín</alternative>
27      <alternative score="0.22222221">César</alternative>
28      <alternative score="0.22222221">Soldado</alternative>
29      <alternative score="0.22222221">Valdés</alternative>
30      <alternative score="0.22222221">Keita</alternative>
31      <alternative score="0.22222221">Griezmann</alternative>
32      <alternative score="0.22222221">Güiza</alternative>
33      <alternative score="0.22222221">Falcao</alternative>
34      <alternative score="0.19999999">Ronaldinho</alternative>
35      <alternative score="0.19999999">Aranzubía</alternative>
36      <alternative score="0.19999999">Rafael Van der Vaart</alternative>
37      <alternative score="0.18181819">Xabi Alonso</alternative>
38      <alternative score="0.18181819">Dani Güiza</alternative>
39      <alternative score="0.18181819">Xavi Prieto</alternative>
40      <alternative score="0.16666669">Sergio Ramos</alternative>
41      <alternative score="0.15384614">Didier Drogba</alternative>
42      <alternative score="0.14285713">Miguel García</alternative>
43      <alternative score="0.14285713">Van Nistelrooy</alternative>
44      <alternative score="0.13333333">Xavi Hernández</alternative>
45      <alternative score="0.13333333">Fernando Torres</alternative>
46      <alternative score="0.11111104">Nolito</alternative>
47      <alternative score="0.11111104">Jeffren</alternative>
48      <alternative score="0.11111104">Iniesta</alternative>
49      <alternative score="0.11111104">Edú</alternative>
50      <alternative score="0.11111104">Raúl</alternative>
51      <alternative score="0.11111104">Palop</alternative>
52      <alternative score="0.11111104">Benzema</alternative>
53      <alternative score="0.11111104">Özil</alternative>
54      <alternative score="0.11111104">Maxwell</alternative>
55      <alternative score="0.11111104">Alves</alternative>
56      <alternative score="0.11111104">Aduriz</alternative>
57      <alternative score="0.11111104">Di María</alternative>
58      <alternative score="0.11111104">Canales</alternative>
59      <alternative score="0.11111104">Cesc</alternative>
60      <alternative score="0.11111104">Riki</alternative>
61      <alternative score="0.11111104">Llorente</alternative>

```

```

62 <alternative score="0.111111104">Pinto</alternative>
63 <alternative score="0.111111104">De Gea</alternative>
64 <alternative score="0.100000024">Dani Alves</alternative>
65 <alternative score="0.090909064">Pedro León</alternative>
66 <alternative score="0.07692307">Miguel Albiol</alternative>
67 <alternative score="0.0">Eto'o</alternative>
68 <alternative score="0.0">Deco</alternative>
69 <alternative score="0.0">Puyol</alternative>
70 <alternative score="0.0">Müller</alternative>
71 <alternative score="0.0">Kun Agüero</alternative>
72 <alternative score="0.0">Leo Messi</alternative>
73 <alternative score="0.0">Lionel Messi</alternative>
74 <alternative score="0.0">Rivaldo</alternative>
75 <alternative score="0.0">Pepe</alternative>
76 <alternative score="0.0">Rosenberg</alternative>
77 <alternative score="0.0">Sneijder</alternative>
78 <alternative score="0.0">Aguirreche</alternative>
79 <alternative score="0.0">Diego Forlán</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.5: Salida para Lass Diarra (voz femenina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="56" end="130" person="UNKNOWN">
4     <correction>
5       <original>Molano</original>
6       <alternative score="0.57142854">Soldado</alternative>
7       <alternative score="0.5">Nolito</alternative>
8       <alternative score="0.39999998">Ronaldinho</alternative>
9       <alternative score="0.3333333">Cristiano</alternative>
10      <alternative score="0.3333333">Tiago</alternative>
11      <alternative score="0.3333333">Falcao</alternative>
12      <alternative score="0.28571427">Forlán</alternative>
13      <alternative score="0.28571427">Rivaldo</alternative>
14      <alternative score="0.27272725">Xabi Alonso</alternative>
15      <alternative score="0.25">Sergio Ramos</alternative>
16      <alternative score="0.25">Carbalho</alternative>
17      <alternative score="0.25">Llorente</alternative>
18      <alternative score="0.22222221">Rosenberg</alternative>
19      <alternative score="0.22222221">Griezmann</alternative>
20      <alternative score="0.19999999">Fernando Torres</alternative>
21      <alternative score="0.17647058">Cristiano Ronaldo</alternative>
22      <alternative score="0.16666669">Eto'o</alternative>
23      <alternative score="0.16666669">Deco</alternative>
24      <alternative score="0.16666669">Xavi</alternative>
25      <alternative score="0.16666669">Villa</alternative>
26      <alternative score="0.16666669">Palop</alternative>
27      <alternative score="0.16666669">Lionel Messi</alternative>
28      <alternative score="0.16666669">Alves</alternative>
29      <alternative score="0.16666669">Pinto</alternative>
30      <alternative score="0.16666669">Carles Puyol</alternative>
31      <alternative score="0.15384614">Andrés Palop</alternative>
32      <alternative score="0.15384614">Miguel Albiol</alternative>
33      <alternative score="0.15384614">Diego Forlán</alternative>
34      <alternative score="0.14999998">Rafael Van der Vaart</alternative>
35      <alternative score="0.14285713">Miguel García</alternative>
36      <alternative score="0.14285713">Müller</alternative>
37      <alternative score="0.14285713">Marcelo</alternative>

```

```

38 <alternative score="0.14285713">Valdés</alternative>
39 <alternative score="0.14285713">Canales</alternative>
40 <alternative score="0.14285713">Van Nistelrooy</alternative>
41 <alternative score="0.125">Casillas</alternative>
42 <alternative score="0.111111104">Leo Messi</alternative>
43 <alternative score="0.111111104">Di María</alternative>
44 <alternative score="0.111111104">Guardiola</alternative>
45 <alternative score="0.100000024">Alcántara</alternative>
46 <alternative score="0.100000024">Dani Alves</alternative>
47 <alternative score="0.100000024">Aranzubía</alternative>
48 <alternative score="0.090909064">Kun Agüero</alternative>
49 <alternative score="0.090909064">Pedro León</alternative>
50 <alternative score="0.090909064">Xavi Prieto</alternative>
51 <alternative score="0.090909064">Lass Diarra</alternative>
52 <alternative score="0.07692307">Gerard Piqué</alternative>
53 <alternative score="0.07692307">Iker Casillas</alternative>
54 <alternative score="0.07692307">Didier Drogba</alternative>
55 <alternative score="0.06666666">Andrés Iniesta</alternative>
56 <alternative score="0.06666666">Xavi Hernández</alternative>
57 <alternative score="0.0">Jeffren</alternative>
58 <alternative score="0.0">Iniesta</alternative>
59 <alternative score="0.0">Puyol</alternative>
60 <alternative score="0.0">David Villa</alternative>
61 <alternative score="0.0">Edú</alternative>
62 <alternative score="0.0">Messi</alternative>
63 <alternative score="0.0">Raúl</alternative>
64 <alternative score="0.0">Benzema</alternative>
65 <alternative score="0.0">Özil</alternative>
66 <alternative score="0.0">Maxwell</alternative>
67 <alternative score="0.0">Santín</alternative>
68 <alternative score="0.0">César</alternative>
69 <alternative score="0.0">Aduriz</alternative>
70 <alternative score="0.0">Pepe</alternative>
71 <alternative score="0.0">Keita</alternative>
72 <alternative score="0.0">Sneijder</alternative>
73 <alternative score="0.0">Cesc</alternative>
74 <alternative score="0.0">Aguirreche</alternative>
75 <alternative score="0.0">Riki</alternative>
76 <alternative score="0.0">Dani Güiza</alternative>
77 <alternative score="0.0">Güiza</alternative>
78 <alternative score="0.0">David De Gea</alternative>
79 <alternative score="0.0">De Gea</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.6: Salida para Müller (voz femenina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="56" end="132" person="UNKNOWN">
4     <correction>
5       <original>0 sea</original>
6       <alternative score="0.3333333">Rosenberg</alternative>
7       <alternative score="0.3333333">De Gea</alternative>
8       <alternative score="0.28571427">Iniesta</alternative>
9       <alternative score="0.28571427">Benzema</alternative>
10      <alternative score="0.28571427">Soldado</alternative>
11      <alternative score="0.25">Casillas</alternative>
12      <alternative score="0.25">Llorente</alternative>
13      <alternative score="0.2222221">Cristiano</alternative>

```

```

14 <alternative score="0.2222221">Leo Messi</alternative>
15 <alternative score="0.2222221">Griezmann</alternative>
16 <alternative score="0.19999999">Villa</alternative>
17 <alternative score="0.19999999">Messi</alternative>
18 <alternative score="0.19999999">Andrés Iniesta</alternative>
19 <alternative score="0.19999999">Alves</alternative>
20 <alternative score="0.19999999">Keita</alternative>
21 <alternative score="0.18181819">Pedro León</alternative>
22 <alternative score="0.18181819">Lass Diarra</alternative>
23 <alternative score="0.16666669">Nolito</alternative>
24 <alternative score="0.16666669">Sergio Ramos</alternative>
25 <alternative score="0.16666669">Lionel Messi</alternative>
26 <alternative score="0.16666669">César</alternative>
27 <alternative score="0.16666669">Güiza</alternative>
28 <alternative score="0.16666669">David De Gea</alternative>
29 <alternative score="0.16666669">Falcao</alternative>
30 <alternative score="0.15384614">Iker Casillas</alternative>
31 <alternative score="0.15384614">Andrés Palop</alternative>
32 <alternative score="0.15384614">Didier Drogba</alternative>
33 <alternative score="0.14285713">Miguel García</alternative>
34 <alternative score="0.14285713">Jeffren</alternative>
35 <alternative score="0.14285713">Forlán</alternative>
36 <alternative score="0.14285713">Müller</alternative>
37 <alternative score="0.14285713">Marcelo</alternative>
38 <alternative score="0.14285713">Maxwell</alternative>
39 <alternative score="0.14285713">Rivaldo</alternative>
40 <alternative score="0.14285713">Canales</alternative>
41 <alternative score="0.14285713">Van Nistelrooy</alternative>
42 <alternative score="0.13333333">Fernando Torres</alternative>
43 <alternative score="0.125">Carvalho</alternative>
44 <alternative score="0.125">Sneijder</alternative>
45 <alternative score="0.11764705">Cristiano Ronaldo</alternative>
46 <alternative score="0.11111104">Di María</alternative>
47 <alternative score="0.11111104">Guardiola</alternative>
48 <alternative score="0.10000024">Ronaldinho</alternative>
49 <alternative score="0.10000024">Alcántara</alternative>
50 <alternative score="0.10000024">Dani Alves</alternative>
51 <alternative score="0.10000024">Aguirreche</alternative>
52 <alternative score="0.10000024">Aranzubia</alternative>
53 <alternative score="0.10000024">Rafael Van der Vaart</alternative>
54 <alternative score="0.090909064">David Villa</alternative>
55 <alternative score="0.090909064">Kun Agüero</alternative>
56 <alternative score="0.090909064">Xabi Alonso</alternative>
57 <alternative score="0.090909064">Dani Güiza</alternative>
58 <alternative score="0.090909064">Xavi Prieto</alternative>
59 <alternative score="0.08333331">Carles Puyol</alternative>
60 <alternative score="0.07692307">Gerard Piqué</alternative>
61 <alternative score="0.07692307">Miguel Albiol</alternative>
62 <alternative score="0.07692307">Diego Forlán</alternative>
63 <alternative score="0.06666666">Xavi Hernández</alternative>
64 <alternative score="0.0">Tiago</alternative>
65 <alternative score="0.0">Eto'o</alternative>
66 <alternative score="0.0">Deco</alternative>
67 <alternative score="0.0">Xavi</alternative>
68 <alternative score="0.0">Puyol</alternative>
69 <alternative score="0.0">Edú</alternative>
70 <alternative score="0.0">Raúl</alternative>
71 <alternative score="0.0">Palop</alternative>
72 <alternative score="0.0">Özil</alternative>
73 <alternative score="0.0">Santín</alternative>
74 <alternative score="0.0">Aduriz</alternative>
75 <alternative score="0.0">Valdés</alternative>

```

```

76     <alternative score="0.0">Pepe</alternative>
77     <alternative score="0.0">Cesc</alternative>
78     <alternative score="0.0">Riki</alternative>
79     <alternative score="0.0">Pinto</alternative>
80   </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.7: Salida para Özil (voz femenina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="69" end="166" person="UNKNOWN">
4     <correction>
5       <original>Rusia no</original>
6       <alternative score="0.666666">Cristiano</alternative>
7       <alternative score="0.4285714">Tiago</alternative>
8       <alternative score="0.35294116">Cristiano Ronaldo</alternative>
9       <alternative score="0.3333333">Griezmann</alternative>
10      <alternative score="0.3076923">Andrés Palop</alternative>
11      <alternative score="0.28571427">Nolito</alternative>
12      <alternative score="0.28571427">Soldado</alternative>
13      <alternative score="0.28571427">Pinto</alternative>
14      <alternative score="0.28571427">Falcao</alternative>
15      <alternative score="0.27272725">Xabi Alonso</alternative>
16      <alternative score="0.27272725">Lass Diarra</alternative>
17      <alternative score="0.25">Casillas</alternative>
18      <alternative score="0.25">Sergio Ramos</alternative>
19      <alternative score="0.25">Carvalho</alternative>
20      <alternative score="0.25">Carles Puyol</alternative>
21      <alternative score="0.23076922">Iker Casillas</alternative>
22      <alternative score="0.22222221">Guardiola</alternative>
23      <alternative score="0.19999999">Ronaldinho</alternative>
24      <alternative score="0.19999999">Andrés Iniesta</alternative>
25      <alternative score="0.19999999">Fernando Torres</alternative>
26      <alternative score="0.18181819">Kun Agüero</alternative>
27      <alternative score="0.18181819">Xavi Prieto</alternative>
28      <alternative score="0.15384614">Gerard Piqué</alternative>
29      <alternative score="0.15384614">Miguel Albiol</alternative>
30      <alternative score="0.15384614">Didier Drogba</alternative>
31      <alternative score="0.14285713">Miguel García</alternative>
32      <alternative score="0.14285713">Eto'o</alternative>
33      <alternative score="0.14285713">Deco</alternative>
34      <alternative score="0.14285713">Xavi</alternative>
35      <alternative score="0.14285713">Villa</alternative>
36      <alternative score="0.14285713">Puyol</alternative>
37      <alternative score="0.14285713">Messi</alternative>
38      <alternative score="0.14285713">Palop</alternative>
39      <alternative score="0.14285713">Marcelo</alternative>
40      <alternative score="0.14285713">Özil</alternative>
41      <alternative score="0.14285713">Rivaldo</alternative>
42      <alternative score="0.14285713">César</alternative>
43      <alternative score="0.14285713">Aduriz</alternative>
44      <alternative score="0.14285713">Keita</alternative>
45      <alternative score="0.14285713">Cesc</alternative>
46      <alternative score="0.14285713">Riki</alternative>
47      <alternative score="0.14285713">Güiza</alternative>
48      <alternative score="0.14285713">Van Nistelrooy</alternative>
49      <alternative score="0.13333333">Xavi Hernández</alternative>
50      <alternative score="0.125">Sneijder</alternative>
51      <alternative score="0.125">Llorente</alternative>

```



```

52 <alternative score="0.111111104">Di María</alternative>
53 <alternative score="0.111111104">Rosenberg</alternative>
54 <alternative score="0.100000024">Alcántara</alternative>
55 <alternative score="0.100000024">Dani Alves</alternative>
56 <alternative score="0.100000024">Aguirreche</alternative>
57 <alternative score="0.100000024">Aranzubia</alternative>
58 <alternative score="0.100000024">Rafael Van der Vaart</alternative>
59 <alternative score="0.090909064">David Villa</alternative>
60 <alternative score="0.090909064">Pedro León</alternative>
61 <alternative score="0.090909064">Dani Güiza</alternative>
62 <alternative score="0.08333331">David De Gea</alternative>
63 <alternative score="0.07692307">Diego Forlán</alternative>
64 <alternative score="0.0">Jeffren</alternative>
65 <alternative score="0.0">Forlán</alternative>
66 <alternative score="0.0">Iniesta</alternative>
67 <alternative score="0.0">Edú</alternative>
68 <alternative score="0.0">Raúl</alternative>
69 <alternative score="0.0">Müller</alternative>
70 <alternative score="0.0">Benzema</alternative>
71 <alternative score="0.0">Leo Messi</alternative>
72 <alternative score="0.0">Lionel Messi</alternative>
73 <alternative score="0.0">Maxwell</alternative>
74 <alternative score="0.0">Alves</alternative>
75 <alternative score="0.0">Santín</alternative>
76 <alternative score="0.0">Valdés</alternative>
77 <alternative score="0.0">Canales</alternative>
78 <alternative score="0.0">Pepe</alternative>
79 <alternative score="0.0">De Gea</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.8: Salida para Rosenberg (voz femenina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="54" end="180" person="UNKNOWN">
4     <correction>
5       <original>C en Girona los</original>
6       <alternative score="0.4166667">Sergio Ramos</alternative>
7       <alternative score="0.3333333">Canales</alternative>
8       <alternative score="0.3076923">Iker Casillas</alternative>
9       <alternative score="0.29411763">Cristiano Ronaldo</alternative>
10      <alternative score="0.26666665">Fernando Torres</alternative>
11      <alternative score="0.25">Casillas</alternative>
12      <alternative score="0.25">Cristiano</alternative>
13      <alternative score="0.25">Tiago</alternative>
14      <alternative score="0.25">Jeffren</alternative>
15      <alternative score="0.25">Ronaldinho</alternative>
16      <alternative score="0.25">Iniesta</alternative>
17      <alternative score="0.25">Palop</alternative>
18      <alternative score="0.25">Benzema</alternative>
19      <alternative score="0.25">Marcelo</alternative>
20      <alternative score="0.25">Carbalho</alternative>
21      <alternative score="0.25">Rivaldo</alternative>
22      <alternative score="0.25">Dani Alves</alternative>
23      <alternative score="0.25">Keita</alternative>
24      <alternative score="0.25">Pinto</alternative>
25      <alternative score="0.23076922">Andrés Palop</alternative>
26      <alternative score="0.21428573">Van Nistelrooy</alternative>
27      <alternative score="0.19999999">Rafael Van der Vaart</alternative>

```



```

28 <alternative score="0.16666669">Nolito</alternative>
29 <alternative score="0.16666669">Eto'o</alternative>
30 <alternative score="0.16666669">Deco</alternative>
31 <alternative score="0.16666669">Villa</alternative>
32 <alternative score="0.16666669">Puyol</alternative>
33 <alternative score="0.16666669">David Villa</alternative>
34 <alternative score="0.16666669">Messi</alternative>
35 <alternative score="0.16666669">Kun Agüero</alternative>
36 <alternative score="0.16666669">Xabi Alonso</alternative>
37 <alternative score="0.16666669">Özil</alternative>
38 <alternative score="0.16666669">Leo Messi</alternative>
39 <alternative score="0.16666669">Santín</alternative>
40 <alternative score="0.16666669">Soldado</alternative>
41 <alternative score="0.16666669">Di María</alternative>
42 <alternative score="0.16666669">Guardiola</alternative>
43 <alternative score="0.16666669">Pedro León</alternative>
44 <alternative score="0.16666669">Sneijder</alternative>
45 <alternative score="0.16666669">Aguirreche</alternative>
46 <alternative score="0.16666669">Llorente</alternative>
47 <alternative score="0.16666669">Griezmann</alternative>
48 <alternative score="0.16666669">Dani Güiza</alternative>
49 <alternative score="0.16666669">Güiza</alternative>
50 <alternative score="0.16666669">De Gea</alternative>
51 <alternative score="0.16666669">Falcao</alternative>
52 <alternative score="0.15384614">Diego Forlán</alternative>
53 <alternative score="0.08333331">Forlán</alternative>
54 <alternative score="0.08333331">Xavi</alternative>
55 <alternative score="0.08333331">Raúl</alternative>
56 <alternative score="0.08333331">Müller</alternative>
57 <alternative score="0.08333331">Lionel Messi</alternative>
58 <alternative score="0.08333331">Maxwell</alternative>
59 <alternative score="0.08333331">Alves</alternative>
60 <alternative score="0.08333331">Alcántara</alternative>
61 <alternative score="0.08333331">César</alternative>
62 <alternative score="0.08333331">Aduriz</alternative>
63 <alternative score="0.08333331">Valdés</alternative>
64 <alternative score="0.08333331">Pepe</alternative>
65 <alternative score="0.08333331">Rosenberg</alternative>
66 <alternative score="0.08333331">Cesc</alternative>
67 <alternative score="0.08333331">Riki</alternative>
68 <alternative score="0.08333331">Aranzubía</alternative>
69 <alternative score="0.08333331">Xavi Prieto</alternative>
70 <alternative score="0.08333331">Carles Puyol</alternative>
71 <alternative score="0.08333331">Lass Diarra</alternative>
72 <alternative score="0.07692307">Gerard Piqué</alternative>
73 <alternative score="0.07692307">Miguel Albiol</alternative>
74 <alternative score="0.07692307">Didier Drogba</alternative>
75 <alternative score="0.0714286">Miguel García</alternative>
76 <alternative score="0.06666666">Andrés Iniesta</alternative>
77 <alternative score="0.06666666">Xavi Hernández</alternative>
78 <alternative score="0.0">Edú</alternative>
79 <alternative score="0.0">David De Gea</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.9: Salida para Sergio Ramos (voz femenina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="62" end="211" person="UNKNOWN">

```

```

4 <correction>
5 <original>Andrés iniesta</original>
6 <alternative score="0.6666666">Andrés Iniesta</alternative>
7 <alternative score="0.5">Iniesta</alternative>
8 <alternative score="0.25">Casillas</alternative>
9 <alternative score="0.25">Benzema</alternative>
10 <alternative score="0.25">Canales</alternative>
11 <alternative score="0.25">Dani Alves</alternative>
12 <alternative score="0.25">Rosenberg</alternative>
13 <alternative score="0.25">Guardiola</alternative>
14 <alternative score="0.25">Keita</alternative>
15 <alternative score="0.25">Pedro León</alternative>
16 <alternative score="0.25">Sneijder</alternative>
17 <alternative score="0.25">Llorente</alternative>
18 <alternative score="0.25">Pinto</alternative>
19 <alternative score="0.25">Dani Güiza</alternative>
20 <alternative score="0.25">Xavi Prieto</alternative>
21 <alternative score="0.25">Lass Diarra</alternative>
22 <alternative score="0.23076922">Andrés Palop</alternative>
23 <alternative score="0.21428573">Van Nistelrooy</alternative>
24 <alternative score="0.19999999">Rafael Van der Vaart</alternative>
25 <alternative score="0.19999999">Fernando Torres</alternative>
26 <alternative score="0.17647058">Cristiano Ronaldo</alternative>
27 <alternative score="0.16666669">Cristiano</alternative>
28 <alternative score="0.16666669">Nolito</alternative>
29 <alternative score="0.16666669">Ronaldinho</alternative>
30 <alternative score="0.16666669">Villa</alternative>
31 <alternative score="0.16666669">David Villa</alternative>
32 <alternative score="0.16666669">Messi</alternative>
33 <alternative score="0.16666669">Marcelo</alternative>
34 <alternative score="0.16666669">Leo Messi</alternative>
35 <alternative score="0.16666669">Lionel Messi</alternative>
36 <alternative score="0.16666669">Alves</alternative>
37 <alternative score="0.16666669">Alcántara</alternative>
38 <alternative score="0.16666669">Santín</alternative>
39 <alternative score="0.16666669">Aduriz</alternative>
40 <alternative score="0.16666669">Valdés</alternative>
41 <alternative score="0.16666669">Cesc</alternative>
42 <alternative score="0.16666669">Aguirreche</alternative>
43 <alternative score="0.16666669">Riki</alternative>
44 <alternative score="0.16666669">Aranzubia</alternative>
45 <alternative score="0.16666669">Griezmann</alternative>
46 <alternative score="0.16666669">Güiza</alternative>
47 <alternative score="0.16666669">David De Gea</alternative>
48 <alternative score="0.16666669">De Gea</alternative>
49 <alternative score="0.15384614">Gerard Piqué</alternative>
50 <alternative score="0.15384614">Iker Casillas</alternative>
51 <alternative score="0.15384614">Didier Drogba</alternative>
52 <alternative score="0.13333333">Xavi Hernández</alternative>
53 <alternative score="0.08333331">Tiago</alternative>
54 <alternative score="0.08333331">Jeffren</alternative>
55 <alternative score="0.08333331">Deco</alternative>
56 <alternative score="0.08333331">Forlán</alternative>
57 <alternative score="0.08333331">Xavi</alternative>
58 <alternative score="0.08333331">Edú</alternative>
59 <alternative score="0.08333331">Müller</alternative>
60 <alternative score="0.08333331">Kun Agüero</alternative>
61 <alternative score="0.08333331">Sergio Ramos</alternative>
62 <alternative score="0.08333331">Carbalho</alternative>
63 <alternative score="0.08333331">Xabi Alonso</alternative>
64 <alternative score="0.08333331">Özil</alternative>
65 <alternative score="0.08333331">Maxwell</alternative>

```

```

66 <alternative score="0.08333331">Rivaldo</alternative>
67 <alternative score="0.08333331">César</alternative>
68 <alternative score="0.08333331">Di María</alternative>
69 <alternative score="0.08333331">Pepe</alternative>
70 <alternative score="0.08333331">Carles Puyol</alternative>
71 <alternative score="0.0714286">Miguel García</alternative>
72 <alternative score="0.0">Eto'o</alternative>
73 <alternative score="0.0">Puyol</alternative>
74 <alternative score="0.0">Raúl</alternative>
75 <alternative score="0.0">Palop</alternative>
76 <alternative score="0.0">Soldado</alternative>
77 <alternative score="0.0">Miguel Albiol</alternative>
78 <alternative score="0.0">Diego Forlán</alternative>
79 <alternative score="0.0">Falcao</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.10: Salida para Andrés Iniesta (voz masculina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="39" end="209" person="UNKNOWN">
4     <correction>
5       <original>Cristiano Ronaldo</original>
6       <alternative score="0.8235294">Cristiano Ronaldo</alternative>
7       <alternative score="0.5">Cristiano</alternative>
8       <alternative score="0.3125">Carvalho</alternative>
9       <alternative score="0.3125">Rivaldo</alternative>
10      <alternative score="0.3125">Griezmann</alternative>
11      <alternative score="0.25">Marcelo</alternative>
12      <alternative score="0.25">Xabi Alonso</alternative>
13      <alternative score="0.25">Soldado</alternative>
14      <alternative score="0.25">Di María</alternative>
15      <alternative score="0.25">Canales</alternative>
16      <alternative score="0.25">Lass Diarra</alternative>
17      <alternative score="0.19999999">Rafael Van der Vaart</alternative>
18      <alternative score="0.1875">Casillas</alternative>
19      <alternative score="0.1875">Tiago</alternative>
20      <alternative score="0.1875">Eto'o</alternative>
21      <alternative score="0.1875">Ronaldinho</alternative>
22      <alternative score="0.1875">Iniesta</alternative>
23      <alternative score="0.1875">Sergio Ramos</alternative>
24      <alternative score="0.1875">Santín</alternative>
25      <alternative score="0.1875">César</alternative>
26      <alternative score="0.1875">Guardiola</alternative>
27      <alternative score="0.1875">Keita</alternative>
28      <alternative score="0.1875">Diego Forlán</alternative>
29      <alternative score="0.1875">Llorente</alternative>
30      <alternative score="0.1875">Aranzubia</alternative>
31      <alternative score="0.1875">Pinto</alternative>
32      <alternative score="0.1875">Xavi Prieto</alternative>
33      <alternative score="0.1875">Didier Drogba</alternative>
34      <alternative score="0.1875">Falcao</alternative>
35      <alternative score="0.125">Nolito</alternative>
36      <alternative score="0.125">Jeffren</alternative>
37      <alternative score="0.125">Forlán</alternative>
38      <alternative score="0.125">Villa</alternative>
39      <alternative score="0.125">Puyol</alternative>
40      <alternative score="0.125">Iker Casillas</alternative>
41      <alternative score="0.125">David Villa</alternative>

```

```

42 <alternative score="0.125">Messi</alternative>
43 <alternative score="0.125">Raúl</alternative>
44 <alternative score="0.125">Andrés Palop</alternative>
45 <alternative score="0.125">Palop</alternative>
46 <alternative score="0.125">Kun Agüero</alternative>
47 <alternative score="0.125">Benzema</alternative>
48 <alternative score="0.125">Özil</alternative>
49 <alternative score="0.125">Lionel Messi</alternative>
50 <alternative score="0.125">Maxwell</alternative>
51 <alternative score="0.125">Alcántara</alternative>
52 <alternative score="0.125">Dani Alves</alternative>
53 <alternative score="0.125">Rosenberg</alternative>
54 <alternative score="0.125">Miguel Albiol</alternative>
55 <alternative score="0.125">Xavi Hernández</alternative>
56 <alternative score="0.125">Sneijder</alternative>
57 <alternative score="0.125">Aguirreche</alternative>
58 <alternative score="0.125">Riki</alternative>
59 <alternative score="0.125">Dani Güiza</alternative>
60 <alternative score="0.125">Güiza</alternative>
61 <alternative score="0.125">Fernando Torres</alternative>
62 <alternative score="0.125">Carles Puyol</alternative>
63 <alternative score="0.125">Van Nistelrooy</alternative>
64 <alternative score="0.0625">Miguel García</alternative>
65 <alternative score="0.0625">Deco</alternative>
66 <alternative score="0.0625">Gerard Piqué</alternative>
67 <alternative score="0.0625">Xavi</alternative>
68 <alternative score="0.0625">Müller</alternative>
69 <alternative score="0.0625">Andrés Iniesta</alternative>
70 <alternative score="0.0625">Leo Messi</alternative>
71 <alternative score="0.0625">Aduriz</alternative>
72 <alternative score="0.0625">Valdés</alternative>
73 <alternative score="0.0625">Pedro León</alternative>
74 <alternative score="0.0625">Cesc</alternative>
75 <alternative score="0.0625">David De Gea</alternative>
76 <alternative score="0.0625">De Gea</alternative>
77 <alternative score="0.0">Edu</alternative>
78 <alternative score="0.0">Alves</alternative>
79 <alternative score="0.0">Pepe</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.11: Salida para Cristiano Ronaldo (voz masculina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="55" end="183" person="UNKNOWN">
4     <correction>
5       <original>David villa</original>
6       <alternative score="0.6363636">David Villa</alternative>
7       <alternative score="0.39999998">Casillas</alternative>
8       <alternative score="0.39999998">Villa</alternative>
9       <alternative score="0.39999998">Guardiola</alternative>
10      <alternative score="0.3333333">David De Gea</alternative>
11      <alternative score="0.3076923">Iker Casillas</alternative>
12      <alternative score="0.3">Carvalho</alternative>
13      <alternative score="0.3">Maxwell</alternative>
14      <alternative score="0.27272725">Xabi Alonso</alternative>
15      <alternative score="0.27272725">Dani Güiza</alternative>
16      <alternative score="0.27272725">Xavi Prieto</alternative>
17      <alternative score="0.27272725">Lass Diarra</alternative>

```

```

18 <alternative score="0.21428573">Van Nistelrooy</alternative>
19 <alternative score="0.19999999">Cristiano</alternative>
20 <alternative score="0.19999999">Xavi</alternative>
21 <alternative score="0.19999999">Iniesta</alternative>
22 <alternative score="0.19999999">Raúl</alternative>
23 <alternative score="0.19999999">Palop</alternative>
24 <alternative score="0.19999999">Marcelo</alternative>
25 <alternative score="0.19999999">Özil</alternative>
26 <alternative score="0.19999999">Andrés Iniesta</alternative>
27 <alternative score="0.19999999">Rivaldo</alternative>
28 <alternative score="0.19999999">Valdés</alternative>
29 <alternative score="0.19999999">Canales</alternative>
30 <alternative score="0.19999999">Dani Alves</alternative>
31 <alternative score="0.19999999">Keita</alternative>
32 <alternative score="0.19999999">Riki</alternative>
33 <alternative score="0.19999999">Aranzubía</alternative>
34 <alternative score="0.19999999">Güiza</alternative>
35 <alternative score="0.19999999">Falcao</alternative>
36 <alternative score="0.15384614">Gerard Piqué</alternative>
37 <alternative score="0.15384614">Andrés Palop</alternative>
38 <alternative score="0.15384614">Miguel Albiol</alternative>
39 <alternative score="0.15384614">Didier Drogba</alternative>
40 <alternative score="0.14999998">Rafael Van der Vaart</alternative>
41 <alternative score="0.13333333">Xavi Hernández</alternative>
42 <alternative score="0.11764705">Cristiano Ronaldo</alternative>
43 <alternative score="0.10000024">Nolito</alternative>
44 <alternative score="0.10000024">Tiago</alternative>
45 <alternative score="0.10000024">Ronaldinho</alternative>
46 <alternative score="0.10000024">Puyol</alternative>
47 <alternative score="0.10000024">Edu</alternative>
48 <alternative score="0.10000024">Messi</alternative>
49 <alternative score="0.10000024">Müller</alternative>
50 <alternative score="0.10000024">Benzema</alternative>
51 <alternative score="0.10000024">Alcántara</alternative>
52 <alternative score="0.10000024">Santín</alternative>
53 <alternative score="0.10000024">Soldado</alternative>
54 <alternative score="0.10000024">Aduriz</alternative>
55 <alternative score="0.10000024">Di María</alternative>
56 <alternative score="0.10000024">Rosenberg</alternative>
57 <alternative score="0.10000024">Sneijder</alternative>
58 <alternative score="0.10000024">Aguirreche</alternative>
59 <alternative score="0.10000024">Griezmann</alternative>
60 <alternative score="0.10000024">Pinto</alternative>
61 <alternative score="0.10000024">De Gea</alternative>
62 <alternative score="0.08333331">Sergio Ramos</alternative>
63 <alternative score="0.08333331">Carles Puyol</alternative>
64 <alternative score="0.07692307">Diego Forlán</alternative>
65 <alternative score="0.0714286">Miguel García</alternative>
66 <alternative score="0.06666666">Fernando Torres</alternative>
67 <alternative score="0.0">Jeffren</alternative>
68 <alternative score="0.0">Eto'o</alternative>
69 <alternative score="0.0">Deco</alternative>
70 <alternative score="0.0">Forlán</alternative>
71 <alternative score="0.0">Kun Agüero</alternative>
72 <alternative score="0.0">Leo Messi</alternative>
73 <alternative score="0.0">Lionel Messi</alternative>
74 <alternative score="0.0">Alves</alternative>
75 <alternative score="0.0">César</alternative>
76 <alternative score="0.0">Pepe</alternative>
77 <alternative score="0.0">Pedro León</alternative>
78 <alternative score="0.0">Cesc</alternative>
79 <alternative score="0.0">Llorente</alternative>

```

```

80     </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.12: Salida para David Villa (voz masculina)

```

1  <?xml version="1.0" encoding="UTF-8" ?>
2  <transcriptions>
3    <sentence id="0" init="47" end="185" person="UNKNOWN">
4      <correction>
5        <original>Iker casillas</original>
6        <alternative score="0.7692308">Iker Casillas</alternative>
7        <alternative score="0.5833334">Casillas</alternative>
8        <alternative score="0.3333333">Villa</alternative>
9        <alternative score="0.3333333">David Villa</alternative>
10       <alternative score="0.3333333">Guardiola</alternative>
11       <alternative score="0.25">Messi</alternative>
12       <alternative score="0.25">Carbalho</alternative>
13       <alternative score="0.25">Canales</alternative>
14       <alternative score="0.25">Keita</alternative>
15       <alternative score="0.25">Falcao</alternative>
16       <alternative score="0.23076922">Andrés Palop</alternative>
17       <alternative score="0.23076922">Didier Drogba</alternative>
18       <alternative score="0.21428573">Miguel García</alternative>
19       <alternative score="0.19999999">Andrés Iniesta</alternative>
20       <alternative score="0.19999999">Fernando Torres</alternative>
21       <alternative score="0.16666669">Cristiano</alternative>
22       <alternative score="0.16666669">Forlán</alternative>
23       <alternative score="0.16666669">Xavi</alternative>
24       <alternative score="0.16666669">Iniesta</alternative>
25       <alternative score="0.16666669">Raúl</alternative>
26       <alternative score="0.16666669">Müller</alternative>
27       <alternative score="0.16666669">Palop</alternative>
28       <alternative score="0.16666669">Sergio Ramos</alternative>
29       <alternative score="0.16666669">Benzema</alternative>
30       <alternative score="0.16666669">Marcelo</alternative>
31       <alternative score="0.16666669">Özil</alternative>
32       <alternative score="0.16666669">Maxwell</alternative>
33       <alternative score="0.16666669">Alves</alternative>
34       <alternative score="0.16666669">Rivaldo</alternative>
35       <alternative score="0.16666669">César</alternative>
36       <alternative score="0.16666669">Aduriz</alternative>
37       <alternative score="0.16666669">Valdés</alternative>
38       <alternative score="0.16666669">Di María</alternative>
39       <alternative score="0.16666669">Dani Alves</alternative>
40       <alternative score="0.16666669">Sneijder</alternative>
41       <alternative score="0.16666669">Cesc</alternative>
42       <alternative score="0.16666669">Riki</alternative>
43       <alternative score="0.16666669">Güiza</alternative>
44       <alternative score="0.16666669">De Gea</alternative>
45       <alternative score="0.15384614">Gerard Piqué</alternative>
46       <alternative score="0.15384614">Miguel Albiol</alternative>
47       <alternative score="0.15384614">Diego Forlán</alternative>
48       <alternative score="0.14999998">Rafael Van der Vaart</alternative>
49       <alternative score="0.14285713">Van Nistelrooy</alternative>
50       <alternative score="0.13333333">Xavi Hernández</alternative>
51       <alternative score="0.11764705">Cristiano Ronaldo</alternative>
52       <alternative score="0.08333331">Nolito</alternative>
53       <alternative score="0.08333331">Tiago</alternative>
54       <alternative score="0.08333331">Jeffren</alternative>
55       <alternative score="0.08333331">Ronaldinho</alternative>

```

```

56 <alternative score="0.08333331">Deco</alternative>
57 <alternative score="0.08333331">Puyol</alternative>
58 <alternative score="0.08333331">Xabi Alonso</alternative>
59 <alternative score="0.08333331">Leo Messi</alternative>
60 <alternative score="0.08333331">Alcántara</alternative>
61 <alternative score="0.08333331">Santín</alternative>
62 <alternative score="0.08333331">Soldado</alternative>
63 <alternative score="0.08333331">Pepe</alternative>
64 <alternative score="0.08333331">Pedro León</alternative>
65 <alternative score="0.08333331">Llorente</alternative>
66 <alternative score="0.08333331">Aranzubía</alternative>
67 <alternative score="0.08333331">Griezmann</alternative>
68 <alternative score="0.08333331">Pinto</alternative>
69 <alternative score="0.08333331">Dani Güiza</alternative>
70 <alternative score="0.08333331">Xavi Prieto</alternative>
71 <alternative score="0.08333331">Carles Puyol</alternative>
72 <alternative score="0.08333331">Lass Diarra</alternative>
73 <alternative score="0.0">Eto'o</alternative>
74 <alternative score="0.0">Edú</alternative>
75 <alternative score="0.0">Kun Agüero</alternative>
76 <alternative score="0.0">Lionel Messi</alternative>
77 <alternative score="0.0">Rosenberg</alternative>
78 <alternative score="0.0">Aguirreche</alternative>
79 <alternative score="0.0">David De Gea</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.13: Salida para Iker Casillas (voz masculina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="59" end="192" person="UNKNOWN">
4     <correction>
5       <original>Lanz Viagra</original>
6       <alternative score="0.45454544">Lass Diarra</alternative>
7       <alternative score="0.3">Tiago</alternative>
8       <alternative score="0.3">Iniesta</alternative>
9       <alternative score="0.3">Benzema</alternative>
10      <alternative score="0.3">Carbalho</alternative>
11      <alternative score="0.3">Alcántara</alternative>
12      <alternative score="0.3">Canales</alternative>
13      <alternative score="0.3">Aranzubía</alternative>
14      <alternative score="0.28571427">Van Nistelrooy</alternative>
15      <alternative score="0.27272725">Dani Güiza</alternative>
16      <alternative score="0.25">Rafael Van der Vaart</alternative>
17      <alternative score="0.19999999">Casillas</alternative>
18      <alternative score="0.19999999">Cristiano</alternative>
19      <alternative score="0.19999999">Xavi</alternative>
20      <alternative score="0.19999999">Villa</alternative>
21      <alternative score="0.19999999">Özil</alternative>
22      <alternative score="0.19999999">Santín</alternative>
23      <alternative score="0.19999999">César</alternative>
24      <alternative score="0.19999999">Dani Alves</alternative>
25      <alternative score="0.19999999">Guardiola</alternative>
26      <alternative score="0.19999999">Keita</alternative>
27      <alternative score="0.19999999">Sneijder</alternative>
28      <alternative score="0.19999999">Güiza</alternative>
29      <alternative score="0.19999999">Fernando Torres</alternative>
30      <alternative score="0.19999999">Falcao</alternative>
31      <alternative score="0.18181819">David Villa</alternative>

```



```

32 <alternative score="0.18181819">Kun Agüero</alternative>
33 <alternative score="0.17647058">Cristiano Ronaldo</alternative>
34 <alternative score="0.16666669">David De Gea</alternative>
35 <alternative score="0.15384614">Gerard Piqué</alternative>
36 <alternative score="0.15384614">Iker Casillas</alternative>
37 <alternative score="0.15384614">Didier Drogba</alternative>
38 <alternative score="0.14285713">Miguel García</alternative>
39 <alternative score="0.13333333">Andrés Iniesta</alternative>
40 <alternative score="0.10000024">Nolito</alternative>
41 <alternative score="0.10000024">Ronaldinho</alternative>
42 <alternative score="0.10000024">Messi</alternative>
43 <alternative score="0.10000024">Raúl</alternative>
44 <alternative score="0.10000024">Müller</alternative>
45 <alternative score="0.10000024">Palop</alternative>
46 <alternative score="0.10000024">Marcelo</alternative>
47 <alternative score="0.10000024">Maxwell</alternative>
48 <alternative score="0.10000024">Rivaldo</alternative>
49 <alternative score="0.10000024">Soldado</alternative>
50 <alternative score="0.10000024">Aduriz</alternative>
51 <alternative score="0.10000024">Valdés</alternative>
52 <alternative score="0.10000024">Di María</alternative>
53 <alternative score="0.10000024">Rosenberg</alternative>
54 <alternative score="0.10000024">Riki</alternative>
55 <alternative score="0.10000024">Griezmann</alternative>
56 <alternative score="0.10000024">Pinto</alternative>
57 <alternative score="0.10000024">De Gea</alternative>
58 <alternative score="0.090909064">Xabi Alonso</alternative>
59 <alternative score="0.090909064">Xavi Prieto</alternative>
60 <alternative score="0.08333331">Sergio Ramos</alternative>
61 <alternative score="0.08333331">Lionel Messi</alternative>
62 <alternative score="0.08333331">Carles Puyol</alternative>
63 <alternative score="0.07692307">Andrés Palop</alternative>
64 <alternative score="0.07692307">Miguel Albiol</alternative>
65 <alternative score="0.07692307">Diego Forlán</alternative>
66 <alternative score="0.06666666">Xavi Hernández</alternative>
67 <alternative score="0.0">Jeffren</alternative>
68 <alternative score="0.0">Eto'o</alternative>
69 <alternative score="0.0">Deco</alternative>
70 <alternative score="0.0">Forlán</alternative>
71 <alternative score="0.0">Puyol</alternative>
72 <alternative score="0.0">Edu</alternative>
73 <alternative score="0.0">Leo Messi</alternative>
74 <alternative score="0.0">Alves</alternative>
75 <alternative score="0.0">Pepe</alternative>
76 <alternative score="0.0">Pedro León</alternative>
77 <alternative score="0.0">Cesc</alternative>
78 <alternative score="0.0">Aguirreche</alternative>
79 <alternative score="0.0">Llorente</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.14: Salida para Lass Diarra (voz masculina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="55" end="159" person="UNKNOWN">
4     <correction>
5       <original>Ni hule</original>
6       <alternative score="0.39999998">Villa</alternative>
7       <alternative score="0.39999998">Dani Alves</alternative>

```



```

8 <alternative score="0.375">Sneijder</alternative>
9 <alternative score="0.28571427">Iniesta</alternative>
10 <alternative score="0.28571427">Müller</alternative>
11 <alternative score="0.28571427">Rivaldo</alternative>
12 <alternative score="0.28571427">Canales</alternative>
13 <alternative score="0.25">Casillas</alternative>
14 <alternative score="0.25">Lionel Messi</alternative>
15 <alternative score="0.23076922">Miguel Albiol</alternative>
16 <alternative score="0.22222221">Guardiola</alternative>
17 <alternative score="0.21428573">Miguel García</alternative>
18 <alternative score="0.21428573">Van Nistelrooy</alternative>
19 <alternative score="0.19999999">Tiago</alternative>
20 <alternative score="0.19999999">Ronaldinho</alternative>
21 <alternative score="0.19999999">Pepe</alternative>
22 <alternative score="0.19999999">Aguirreche</alternative>
23 <alternative score="0.19999999">Riki</alternative>
24 <alternative score="0.19999999">Aranzubía</alternative>
25 <alternative score="0.19999999">Pinto</alternative>
26 <alternative score="0.18181819">David Villa</alternative>
27 <alternative score="0.18181819">Kun Agüero</alternative>
28 <alternative score="0.18181819">Xabi Alonso</alternative>
29 <alternative score="0.18181819">Dani Güiza</alternative>
30 <alternative score="0.18181819">Xavi Prieto</alternative>
31 <alternative score="0.16666669">Aduriz</alternative>
32 <alternative score="0.16666669">David De Gea</alternative>
33 <alternative score="0.16666669">De Gea</alternative>
34 <alternative score="0.16666669">Carles Puyol</alternative>
35 <alternative score="0.15384614">Gerard Piqué</alternative>
36 <alternative score="0.15384614">Iker Casillas</alternative>
37 <alternative score="0.15384614">Andrés Palop</alternative>
38 <alternative score="0.15384614">Diego Forlán</alternative>
39 <alternative score="0.15384614">Didier Drogba</alternative>
40 <alternative score="0.14285713">Jeffren</alternative>
41 <alternative score="0.14285713">Forlán</alternative>
42 <alternative score="0.14285713">Benzema</alternative>
43 <alternative score="0.14285713">Marcelo</alternative>
44 <alternative score="0.14285713">Maxwell</alternative>
45 <alternative score="0.14285713">Santín</alternative>
46 <alternative score="0.13333333">Andrés Iniesta</alternative>
47 <alternative score="0.13333333">Xavi Hernández</alternative>
48 <alternative score="0.13333333">Fernando Torres</alternative>
49 <alternative score="0.125">Carbalho</alternative>
50 <alternative score="0.125">Llorente</alternative>
51 <alternative score="0.11764705">Cristiano Ronaldo</alternative>
52 <alternative score="0.11111104">Cristiano</alternative>
53 <alternative score="0.11111104">Leo Messi</alternative>
54 <alternative score="0.11111104">Di María</alternative>
55 <alternative score="0.11111104">Rosenberg</alternative>
56 <alternative score="0.11111104">Griezmann</alternative>
57 <alternative score="0.10000024">Alcántara</alternative>
58 <alternative score="0.10000024">Rafael Van der Vaart</alternative>
59 <alternative score="0.090909064">Pedro León</alternative>
60 <alternative score="0.090909064">Lass Diarra</alternative>
61 <alternative score="0.08333331">Sergio Ramos</alternative>
62 <alternative score="0.0">Nolito</alternative>
63 <alternative score="0.0">Eto'o</alternative>
64 <alternative score="0.0">Deco</alternative>
65 <alternative score="0.0">Xavi</alternative>
66 <alternative score="0.0">Puyol</alternative>
67 <alternative score="0.0">Edú</alternative>
68 <alternative score="0.0">Messi</alternative>
69 <alternative score="0.0">Raúl</alternative>

```

```

70 <alternative score="0.0">Palop</alternative>
71 <alternative score="0.0">Özil</alternative>
72 <alternative score="0.0">Alves</alternative>
73 <alternative score="0.0">César</alternative>
74 <alternative score="0.0">Soldado</alternative>
75 <alternative score="0.0">Valdés</alternative>
76 <alternative score="0.0">Keita</alternative>
77 <alternative score="0.0">Cesc</alternative>
78 <alternative score="0.0">Güiza</alternative>
79 <alternative score="0.0">Falcao</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.15: Salida para Müller (voz masculina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="39" end="124" person="UNKNOWN">
4     <correction>
5       <original>0 civil</original>
6       <alternative score="0.333333">Özil</alternative>
7       <alternative score="0.333333">Riki</alternative>
8       <alternative score="0.27272725">David Villa</alternative>
9       <alternative score="0.25">Carbalho</alternative>
10      <alternative score="0.23076922">Miguel Albiol</alternative>
11      <alternative score="0.22222221">Cristiano</alternative>
12      <alternative score="0.22222221">Rosenberg</alternative>
13      <alternative score="0.22222221">Guardiola</alternative>
14      <alternative score="0.19999999">Ronaldo</alternative>
15      <alternative score="0.19999999">Dani Alves</alternative>
16      <alternative score="0.19999999">Aranzubía</alternative>
17      <alternative score="0.18181819">Xabi Alonso</alternative>
18      <alternative score="0.18181819">Dani Güiza</alternative>
19      <alternative score="0.18181819">Xavi Prieto</alternative>
20      <alternative score="0.17647058">Cristiano Ronaldo</alternative>
21      <alternative score="0.16666669">Nolito</alternative>
22      <alternative score="0.16666669">Tiago</alternative>
23      <alternative score="0.16666669">Xavi</alternative>
24      <alternative score="0.16666669">Villa</alternative>
25      <alternative score="0.16666669">Puyol</alternative>
26      <alternative score="0.16666669">Messi</alternative>
27      <alternative score="0.16666669">Raúl</alternative>
28      <alternative score="0.16666669">Aduriz</alternative>
29      <alternative score="0.16666669">Keita</alternative>
30      <alternative score="0.16666669">Pinto</alternative>
31      <alternative score="0.15384614">Iker Casillas</alternative>
32      <alternative score="0.15384614">Diego Forlán</alternative>
33      <alternative score="0.14285713">Forlán</alternative>
34      <alternative score="0.14285713">Iniesta</alternative>
35      <alternative score="0.14285713">Marcelo</alternative>
36      <alternative score="0.14285713">Maxwell</alternative>
37      <alternative score="0.14285713">Rivaldo</alternative>
38      <alternative score="0.14285713">Soldado</alternative>
39      <alternative score="0.14285713">Van Nistelrooy</alternative>
40      <alternative score="0.125">Casillas</alternative>
41      <alternative score="0.125">Sneijder</alternative>
42      <alternative score="0.125">Llorente</alternative>
43      <alternative score="0.111111104">Leo Messi</alternative>
44      <alternative score="0.111111104">Griezmann</alternative>
45      <alternative score="0.100000024">Aguirreche</alternative>

```

```

46 <alternative score="0.090909064">Pedro León</alternative>
47 <alternative score="0.090909064">Lass Diarra</alternative>
48 <alternative score="0.08333331">Sergio Ramos</alternative>
49 <alternative score="0.08333331">Lionel Messi</alternative>
50 <alternative score="0.08333331">David De Gea</alternative>
51 <alternative score="0.08333331">Carles Puyol</alternative>
52 <alternative score="0.07692307">Gerard Piqué</alternative>
53 <alternative score="0.07692307">Andrés Palop</alternative>
54 <alternative score="0.07692307">Didier Drogba</alternative>
55 <alternative score="0.0714286">Miguel García</alternative>
56 <alternative score="0.06666666">Andrés Iniesta</alternative>
57 <alternative score="0.06666666">Xavi Hernández</alternative>
58 <alternative score="0.06666666">Fernando Torres</alternative>
59 <alternative score="0.05000012">Rafael Van der Vaart</alternative>
60 <alternative score="0.0">Jeffren</alternative>
61 <alternative score="0.0">Eto'o</alternative>
62 <alternative score="0.0">Deco</alternative>
63 <alternative score="0.0">Edú</alternative>
64 <alternative score="0.0">Müller</alternative>
65 <alternative score="0.0">Palop</alternative>
66 <alternative score="0.0">Kun Agüero</alternative>
67 <alternative score="0.0">Benzema</alternative>
68 <alternative score="0.0">Alves</alternative>
69 <alternative score="0.0">Alcántara</alternative>
70 <alternative score="0.0">Santín</alternative>
71 <alternative score="0.0">César</alternative>
72 <alternative score="0.0">Valdés</alternative>
73 <alternative score="0.0">Di María</alternative>
74 <alternative score="0.0">Canales</alternative>
75 <alternative score="0.0">Pepe</alternative>
76 <alternative score="0.0">Cesc</alternative>
77 <alternative score="0.0">Güiza</alternative>
78 <alternative score="0.0">De Gea</alternative>
79 <alternative score="0.0">Falcao</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.16: Salida para Özil (voz masculina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="60" end="160" person="UNKNOWN">
4     <correction>
5       <original>Rozan de</original>
6       <alternative score="0.375">Llorente</alternative>
7       <alternative score="0.3333333">Rosenberg</alternative>
8       <alternative score="0.3333333">Griezmann</alternative>
9       <alternative score="0.3">Ronaldinho</alternative>
10      <alternative score="0.28571427">Rivaldo</alternative>
11      <alternative score="0.28571427">Soldado</alternative>
12      <alternative score="0.26666665">Xavi Hernández</alternative>
13      <alternative score="0.26666665">Fernando Torres</alternative>
14      <alternative score="0.25">Sergio Ramos</alternative>
15      <alternative score="0.25">Sneijder</alternative>
16      <alternative score="0.2352941">Cristiano Ronaldo</alternative>
17      <alternative score="0.22222221">Cristiano</alternative>
18      <alternative score="0.19999999">Andrés Iniesta</alternative>
19      <alternative score="0.19999999">Aranzubía</alternative>
20      <alternative score="0.19999999">Rafael Van der Vaart</alternative>
21      <alternative score="0.18181819">Pedro León</alternative>

```

```

22 <alternative score="0.16666669">Lionel Messi</alternative>
23 <alternative score="0.15384614">Gerard Piqué</alternative>
24 <alternative score="0.15384614">Iker Casillas</alternative>
25 <alternative score="0.15384614">Andrés Palop</alternative>
26 <alternative score="0.14285713">Nolito</alternative>
27 <alternative score="0.14285713">Tiago</alternative>
28 <alternative score="0.14285713">Forlán</alternative>
29 <alternative score="0.14285713">Xavi</alternative>
30 <alternative score="0.14285713">Raúl</alternative>
31 <alternative score="0.14285713">Palop</alternative>
32 <alternative score="0.14285713">Özil</alternative>
33 <alternative score="0.14285713">Canales</alternative>
34 <alternative score="0.14285713">Pepe</alternative>
35 <alternative score="0.14285713">Pinto</alternative>
36 <alternative score="0.125">Carvalho</alternative>
37 <alternative score="0.11111104">Leo Messi</alternative>
38 <alternative score="0.11111104">Di María</alternative>
39 <alternative score="0.11111104">Guardiola</alternative>
40 <alternative score="0.10000024">Alcántara</alternative>
41 <alternative score="0.10000024">Dani Alves</alternative>
42 <alternative score="0.10000024">Aguirreche</alternative>
43 <alternative score="0.090909064">Kun Agüero</alternative>
44 <alternative score="0.090909064">Xabi Alonso</alternative>
45 <alternative score="0.090909064">Xavi Prieto</alternative>
46 <alternative score="0.090909064">Lass Diarra</alternative>
47 <alternative score="0.08333331">David De Gea</alternative>
48 <alternative score="0.08333331">Carles Puyol</alternative>
49 <alternative score="0.07692307">Diego Forlán</alternative>
50 <alternative score="0.07692307">Didier Drogba</alternative>
51 <alternative score="0.0714286">Miguel García</alternative>
52 <alternative score="0.0714286">Van Nistelrooy</alternative>
53 <alternative score="0.0">Casillas</alternative>
54 <alternative score="0.0">Jeffren</alternative>
55 <alternative score="0.0">Eto'o</alternative>
56 <alternative score="0.0">Deco</alternative>
57 <alternative score="0.0">Villa</alternative>
58 <alternative score="0.0">Iniesta</alternative>
59 <alternative score="0.0">Puyol</alternative>
60 <alternative score="0.0">David Villa</alternative>
61 <alternative score="0.0">Edú</alternative>
62 <alternative score="0.0">Messi</alternative>
63 <alternative score="0.0">Müller</alternative>
64 <alternative score="0.0">Benzema</alternative>
65 <alternative score="0.0">Marcelo</alternative>
66 <alternative score="0.0">Maxwell</alternative>
67 <alternative score="0.0">Alves</alternative>
68 <alternative score="0.0">Santín</alternative>
69 <alternative score="0.0">César</alternative>
70 <alternative score="0.0">Aduriz</alternative>
71 <alternative score="0.0">Valdés</alternative>
72 <alternative score="0.0">Miguel Albiol</alternative>
73 <alternative score="0.0">Keita</alternative>
74 <alternative score="0.0">Cesc</alternative>
75 <alternative score="0.0">Riki</alternative>
76 <alternative score="0.0">Dani Güiza</alternative>
77 <alternative score="0.0">Güiza</alternative>
78 <alternative score="0.0">De Gea</alternative>
79 <alternative score="0.0">Falcao</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.17: Salida para Rosenberg (voz masculina)

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <transcriptions>
3   <sentence id="0" init="53" end="187" person="UNKNOWN">
4     <correction>
5       <original>Sergio ramos</original>
6       <alternative score="0.6666666">Sergio Ramos</alternative>
7       <alternative score="0.27272725">Casillas</alternative>
8       <alternative score="0.27272725">Cristiano</alternative>
9       <alternative score="0.27272725">Tiago</alternative>
10      <alternative score="0.27272725">Guardiola</alternative>
11      <alternative score="0.27272725">Keita</alternative>
12      <alternative score="0.26666665">Fernando Torres</alternative>
13      <alternative score="0.23076922">Iker Casillas</alternative>
14      <alternative score="0.23076922">Andrés Palop</alternative>
15      <alternative score="0.23076922">Diego Forlán</alternative>
16      <alternative score="0.18181819">Nolito</alternative>
17      <alternative score="0.18181819">Jeffren</alternative>
18      <alternative score="0.18181819">Eto'o</alternative>
19      <alternative score="0.18181819">Deco</alternative>
20      <alternative score="0.18181819">Forlán</alternative>
21      <alternative score="0.18181819">Villa</alternative>
22      <alternative score="0.18181819">Messi</alternative>
23      <alternative score="0.18181819">Palop</alternative>
24      <alternative score="0.18181819">Benzema</alternative>
25      <alternative score="0.18181819">Marcelo</alternative>
26      <alternative score="0.18181819">Carbalho</alternative>
27      <alternative score="0.18181819">Leo Messi</alternative>
28      <alternative score="0.18181819">Rivaldo</alternative>
29      <alternative score="0.18181819">Soldado</alternative>
30      <alternative score="0.18181819">Canales</alternative>
31      <alternative score="0.18181819">Dani Alves</alternative>
32      <alternative score="0.18181819">Sneijder</alternative>
33      <alternative score="0.18181819">Aguirreche</alternative>
34      <alternative score="0.18181819">Llorente</alternative>
35      <alternative score="0.18181819">Griezmann</alternative>
36      <alternative score="0.18181819">Pinto</alternative>
37      <alternative score="0.18181819">Güiza</alternative>
38      <alternative score="0.18181819">De Gea</alternative>
39      <alternative score="0.18181819">Falcao</alternative>
40      <alternative score="0.17647058">Cristiano Ronaldo</alternative>
41      <alternative score="0.16666669">Carles Puyol</alternative>
42      <alternative score="0.15384614">Gerard Piqué</alternative>
43      <alternative score="0.15384614">Didier Drogba</alternative>
44      <alternative score="0.14999998">Rafael Van der Vaart</alternative>
45      <alternative score="0.14285713">Miguel García</alternative>
46      <alternative score="0.14285713">Van Nistelrooy</alternative>
47      <alternative score="0.13333333">Andrés Iniesta</alternative>
48      <alternative score="0.090909064">Ronaldinho</alternative>
49      <alternative score="0.090909064">Xavi</alternative>
50      <alternative score="0.090909064">Iniesta</alternative>
51      <alternative score="0.090909064">Puyol</alternative>
52      <alternative score="0.090909064">Raúl</alternative>
53      <alternative score="0.090909064">Müller</alternative>
54      <alternative score="0.090909064">Xabi Alonso</alternative>
55      <alternative score="0.090909064">Özil</alternative>
56      <alternative score="0.090909064">Alves</alternative>
57      <alternative score="0.090909064">Alcántara</alternative>

```

```

58 <alternative score="0.090909064">César</alternative>
59 <alternative score="0.090909064">Aduriz</alternative>
60 <alternative score="0.090909064">Valdés</alternative>
61 <alternative score="0.090909064">Di María</alternative>
62 <alternative score="0.090909064">Pepe</alternative>
63 <alternative score="0.090909064">Rosenberg</alternative>
64 <alternative score="0.090909064">Pedro León</alternative>
65 <alternative score="0.090909064">Cesc</alternative>
66 <alternative score="0.090909064">Riki</alternative>
67 <alternative score="0.090909064">Aranzubía</alternative>
68 <alternative score="0.090909064">Xavi Prieto</alternative>
69 <alternative score="0.090909064">Lass Diarra</alternative>
70 <alternative score="0.08333331">Lionel Messi</alternative>
71 <alternative score="0.07692307">Miguel Albiol</alternative>
72 <alternative score="0.06666666">Xavi Hernández</alternative>
73 <alternative score="0.0">David Villa</alternative>
74 <alternative score="0.0">Edú</alternative>
75 <alternative score="0.0">Kun Agüero</alternative>
76 <alternative score="0.0">Maxwell</alternative>
77 <alternative score="0.0">Santín</alternative>
78 <alternative score="0.0">Dani Güiza</alternative>
79 <alternative score="0.0">David De Gea</alternative>
80 </correction>
81 </sentence>
82 </transcriptions>

```

Listing E.18: Salida para Sergio Ramos (voz masculina)

E.2. Resultado de la evaluación

Estos son los ficheros utilizados para la evaluación con sclite y los resultados devueltos:

```

1 Andrés Iniesta (s_01)
2 Cristiano Ronaldo (s_02)
3 David Villa (s_03)
4 Iker Casillas (s_04)
5 Lass Diarra (s_05)
6 Müller (s_06)
7 Özil (s_07)
8 Rosenberg (s_08)
9 Sergio Ramos (s_09)

```

Listing E.19: TRN de referencia para Sc-lite

```

1 Andrés iniesta (s_01)
2 Cristianos Donaldo (s_02)
3 David villa (s_03)
4 Iker casillas (s_04)
5 Lars diana (s_05)
6 Molano (s_06)
7 O sea (s_07)
8 Rusia no (s_08)
9 C en Girona los (s_09)

```

Listing E.20: TRN de voz femenina sin corregir

[illegible]

```

62 | -----
63 |         0
64 |
65 |
66 |
67 | SUBSTITUTIONS                                Total                (9)
68 |                                         With >= 1 occurrences (9)
69 |
70 | 1:    1  -> cristiano
71 | 2:    1  -> diarra
72 | 3:    1  -> lass
73 | 4:    1  -> müller
74 | 5:    1  -> ramos
75 | 6:    1  -> ronaldo
76 | 7:    1  -> rosenberg
77 | 8:    1  -> sergio
78 | 9:    1  -> özil
79 | -----
80 |         9
81 |
82 |
83 | * NOTE: The 'Substitution' words are those reference words
84 |         for which the recognizer supplied an incorrect word.
85 |
86 |
87 | FALSELY RECOGNIZED                        Total                (9)
88 |                                         With >= 1 occurrences (9)
89 |
90 | 1:    1  -> cristianos
91 | 2:    1  -> diana
92 | 3:    1  -> donaldo
93 | 4:    1  -> girona
94 | 5:    1  -> lars
95 | 6:    1  -> los
96 | 7:    1  -> molano
97 | 8:    1  -> no
98 | 9:    1  -> sea
99 | -----
100 |        9
101 |
102 |
103 | * NOTE: The 'Falsely Recognized' words are those hypothesis words
104 |         which the recognizer incorrectly substituted for a reference word.

```

Listing E.21: Resultado para voz femenina sin corregir

```

1 | Andrés Iniesta (s_01)
2 | Cristiano Ronaldo (s_02)
3 | David Villa (s_03)
4 | Iker Casillas (s_04)
5 | Lass Diarra (s_05)
6 | Soldado (s_06)
7 | Rosenberg (s_07)
8 | Cristiano (s_08)
9 | Sergio Ramos (s_09)

```

Listing E.22: TRN de voz femenina corregido

```

1 | SCORING FOR SPEAKER: s_0
2 |   of f_correct.trn

```



```

3
4 SENTENCE RECOGNITION PERFORMANCE
5
6 sentences                      9
7 with errors                    33.3% ( 3)
8
9   with substitutions           33.3% ( 3)
10  with deletions               0.0% ( 0)
11  with insertions              0.0% ( 0)
12
13
14 WORD RECOGNITION PERFORMANCE
15
16 Percent Total Error           = 20.0% ( 3)
17
18 Percent Correct               = 80.0% ( 12)
19
20 Percent Substitution          = 20.0% ( 3)
21 Percent Deletions             = 0.0% ( 0)
22 Percent Insertions           = 0.0% ( 0)
23 Percent Word Accuracy         = 80.0%
24
25
26 Ref. words                    = ( 15)
27 Hyp. words                    = ( 15)
28 Aligned words                 = ( 15)
29
30 CONFUSION PAIRS                Total (3)
31                               With >= 1 occurrences (3)
32
33 1: 1 -> müller ==> soldado
34 2: 1 -> rosenberg ==> cristiano
35 3: 1 -> özil ==> rosenberg
36 -----
37      3
38
39
40
41 INSERTIONS                    Total (0)
42                               With >= 1 occurrences (0)
43
44 -----
45      0
46
47
48
49 DELETIONS                     Total (0)
50                               With >= 1 occurrences (0)
51
52 -----
53      0
54
55
56
57 SUBSTITUTIONS                 Total (3)
58                               With >= 1 occurrences (3)
59
60 1: 1 -> müller
61 2: 1 -> rosenberg
62 3: 1 -> özil
63 -----
64      3

```

```

65
66
67 * NOTE: The 'Substitution' words are those reference words
68       for which the recognizer supplied an incorrect word.
69
70
71 FALSELY RECOGNIZED                Total                (3)
72                                With >= 1 occurrences (3)
73
74     1:      1 -> cristiano
75     2:      1 -> rosenberg
76     3:      1 -> soldado
77     -----
78           3
79
80
81 * NOTE: The 'Falsely Recognized' words are those hypothesis words
82       which the recognizer incorrectly substituted for a reference word.

```

Listing E.23: Resultado para voz femenina corregida

```

1 Andrés iniesta (s_01)
2 Cristiano Ronaldo (s_02)
3 David villa (s_03)
4 Iker casillas (s_04)
5 Lanz Viagra (s_05)
6 Ni hule (s_06)
7 O civil (s_07)
8 Rozan de (s_08)
9 Sergio ramos (s_09)

```

Listing E.24: TRN de voz masculino sin corregir

```

1 SCORING FOR SPEAKER: s_0
2   of m_raw.trn
3
4 SENTENCE RECOGNITION PERFORMANCE
5
6 sentences                                9
7 with errors                            44.4% ( 4)
8
9   with substitutions                    44.4% ( 4)
10  with deletions                        0.0% ( 0)
11  with insertions                       33.3% ( 3)
12
13
14 WORD RECOGNITION PERFORMANCE
15
16 Percent Total Error      = 53.3% ( 8)
17
18 Percent Correct          = 66.7% ( 10)
19
20 Percent Substitution      = 33.3% ( 5)
21 Percent Deletions         = 0.0% ( 0)
22 Percent Insertions        = 20.0% ( 3)
23 Percent Word Accuracy     = 46.7%
24
25
26 Ref. words                = ( 15)
27 Hyp. words                 = ( 18)

```

```

28 Aligned words          =          ( 18)
29
30 CONFUSION PAIRS          Total          (5)
31          With >= 1 occurances (5)
32
33 1: 1 -> diarra ==> viagra
34 2: 1 -> lass ==> lanz
35 3: 1 -> müller ==> hule
36 4: 1 -> rosenberg ==> de
37 5: 1 -> özil ==> civil
38 -----
39 5
40
41
42
43 INSERTIONS          Total          (3)
44          With >= 1 occurances (3)
45
46 1: 1 -> ni
47 2: 1 -> o
48 3: 1 -> rozan
49 -----
50 3
51
52
53
54 DELETIONS          Total          (0)
55          With >= 1 occurances (0)
56
57 -----
58 0
59
60
61
62 SUBSTITUTIONS          Total          (5)
63          With >= 1 occurances (5)
64
65 1: 1 -> diarra
66 2: 1 -> lass
67 3: 1 -> müller
68 4: 1 -> rosenberg
69 5: 1 -> özil
70 -----
71 5
72
73
74 * NOTE: The 'Substitution' words are those reference words
75         for which the recognizer supplied an incorrect word.
76
77
78 FALSELY RECOGNIZED          Total          (5)
79          With >= 1 occurances (5)
80
81 1: 1 -> civil
82 2: 1 -> de
83 3: 1 -> hule
84 4: 1 -> lanz
85 5: 1 -> viagra
86 -----
87 5
88
89

```

90 * NOTE: The 'Falsely Recognized' words are those hypothesis words
 91 which the recognizer incorrectly substituted for a reference word.

Listing E.25: Resultado para voz masculina sin corregir

```

1 Andrés Iniesta (s_01)
2 Cristiano Ronaldo (s_02)
3 David Villa (s_03)
4 Iker Casillas (s_04)
5 Lass Diarra (s_05)
6 Villa (s_06)
7 Özil (s_07)
8 Llorente (s_08)
9 Sergio Ramos (s_09)

```

Listing E.26: TRN de voz masculino corregido

```

1 SCORING FOR SPEAKER: s_0
2   of m_correct.trn
3
4 SENTENCE RECOGNITION PERFORMANCE
5
6 sentences                      9
7 with errors                    22.2% ( 2)
8
9   with substitutions           22.2% ( 2)
10  with deletions                0.0% ( 0)
11  with insertions              0.0% ( 0)
12
13
14 WORD RECOGNITION PERFORMANCE
15
16 Percent Total Error           = 13.3% ( 2)
17
18 Percent Correct               = 86.7% ( 13)
19
20 Percent Substitution          = 13.3% ( 2)
21 Percent Deletions             = 0.0% ( 0)
22 Percent Insertions           = 0.0% ( 0)
23 Percent Word Accuracy         = 86.7%
24
25
26 Ref. words                    = ( 15)
27 Hyp. words                    = ( 15)
28 Aligned words                 = ( 15)
29
30 CONFUSION PAIRS                      Total (2)
31                                     With >= 1 occurrences (2)
32
33 1: 1 -> müller ==> villa
34 2: 1 -> rosenberg ==> llorente
35 -----
36      2
37
38
39
40 INSERTIONS                      Total (0)
41                                     With >= 1 occurrences (0)
42 -----
43

```

```

44      0
45
46
47
48 DELETIONS                                Total                (0)
49                                         With >= 1 occurances (0)
50
51      -----
52      0
53
54
55
56 SUBSTITUTIONS                            Total                (2)
57                                         With >= 1 occurances (2)
58
59      1:    1  -> müller
60      2:    1  -> rosenberg
61      -----
62      2
63
64
65 * NOTE: The 'Substitution' words are those reference words
66         for which the recognizer supplied an incorrect word.
67
68
69 FALSELY RECOGNIZED                       Total                (2)
70                                         With >= 1 occurances (2)
71
72      1:    1  -> llorente
73      2:    1  -> villa
74      -----
75      2
76
77
78 * NOTE: The 'Falsely Recognized' words are those hypothesis words
79         which the recognizer incorrectly substituted for a reference word.

```

Listing E.27: Resultado para voz masculina corregida

Índice alfabético

API, [3](#), [37](#)
asistente personal, [5](#)
Dragon Naturally Speaking, [13](#)
hipervisor, [16](#)
identificación biométrica, [8](#)
Nina, [8](#)
Nuance, [8](#)
S Voice, [7](#)
SaaS, [4](#)
Samsung, [7](#)
servicio web, [8](#), [13](#)
Sherpa, [8](#)
Siri, [5](#)
Windows Speech API, [13](#), [39](#)
Xen, [16](#)